(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 August 2001 (30.08.2001)

PCT

(10) International Publication Number
WO 01/62794 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/05529

(22) International Filing Date: 20 February 2001 (20.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/510,706	22 February 2000 (22.02.2000)	US
09/583,373	31 May 2000 (31.05.2000)	US
09/634,669	8 August 2000 (08.08.2000)	US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GLUCKSMANN, Maria, Alexandra** [AR/US]; 33 Summit Road, Lexington, MA 02173 (US). **CURTIS, Rory, A., J.** [US/US]; 31 Constitution Drive, Southborough, MA 01772 (US). **LORA, Jose, M.** [US/US]; 49 North Main Street, Natick, MA 01760 (US).

(74) Agents: **MANDRAGOURAS, Amy, E.**; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/62794 A2

(54) Title: 18607, A NOVEL HUMAN CALCIUM CHANNEL

(57) Abstract: The invention provides isolated nucleic acids molecules, designated TLCC nucleic acid molecules, which encode novel TRP-like calcium channel molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TLCC nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TLCC gene has been introduced or disrupted. The invention still further provides isolated TLCC proteins, fusion proteins, antigenic peptides and anti-TLCC antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

18607, A NOVEL HUMAN CALCIUM CHANNEL

Related Applications

This application claims priority to U.S. Patent Application No.: 09/634,669, filed on August 8, 2000, U.S. Patent Application No.: 09/583,373, filed on May 31, 2000, and U.S. Patent Application No.: 09/510,706, filed on February 22, 2000, incorporated herein in their entirety by this reference.

Background of the Invention

Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane, or calcium ions may be freed from intracellular stores, again being transported by specific membrane channels in the storage organelle. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (for review, see Putney and McKay (1999) *BioEssays* 21:38-46).

Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated nonselective cation current (CRANC) (Krause *et al.* (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins. There is no single electrophysiological profile characteristic of the family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different specific channels. Further, in several instances it has been demonstrated that homo- or heteropolymerization of the channel molecule may occur, further changing the channel properties from that of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated upon stimulation of phospholipase C_β by a G protein-coupled receptor. Depletion of intracellular calcium stores activate these channels by a

mechanism which is as yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (*e.g.*, by the introduction of chelators into the cell, by activating phospholipase C_γ, or by inhibiting the those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion concentrations) (Putney, J.W., (1986) *Cell Calcium* 7: 1-12; Putney, J.W. (1990) *Cell Calcium* 11:611-624).

The TRP channel family is one of the best characterized of the capacitative calcium channel group. These channels include transient receptor potential protein and homologues thereof (to date, seven homologs and splice variants have been identified in a variety of organisms), the vanilloid receptor subtype I (also known as the capsaicin receptor), stretch-inhibitable non-selective cation channel (SIC), olfactory, mechanosensitive channel, insulin-like growth factor I-regulated calcium channel, and vitamin D-responsive apical, epithelial calcium channel (ECaC) (see, *e.g.*, Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina *et al.* (1997) *Nature* 389: 816-824; Suzuki *et al.* (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov *et al.* (1998) *Nature* 396: 478-482; and Hoenderop *et al.* (1999) *J. Biol. Chem.* 274: 8375-8378). Each of these molecules is 700 or more amino acids (TRP and TRP homologs have 1300 or more amino acid residues), and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains and frequently display a proline-rich region at the N-terminus. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain (see, *e.g.*, McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light (Hardie and Minke, *supra*), or olfactory signals (Colbert *et al.* (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

Calcium signaling may play a role in liver disease. Ca^{2+} influx has been shown to be essential for the contractile phenotype of activated stellate cells, being the phenotype considered responsible for the high portal hypertension associated with hepatic fibrosis. Hepatic stellate cells, a scarce liver cell type, have been proposed as the main effector of the fibrotic process. Once stimulated, stellate cells acquire the activated phenotype, proliferate, and become fibrogenic. Activated stellate cells contribute to the build-up of extracellular matrix (ECM) via overproduction of ECM components (*e.g.*, collagen), and inhibition of their breakdown. The stimuli for stellate cell activation are not yet clear, although inflammatory cells (*e.g.*, T-lymphocytes) and their mediators (*e.g.*, growth factors, cytokines, and chemokines) interacting with their specific receptors (*e.g.*, GPCRs), have all been postulated to play a role. In addition, PDGF-mediated stellate cell proliferation (a key phenotype of activated stellate cells) depends on Ca^{2+} influx.

Vascular Disorders

Cardiovascular disease is a major health risk throughout the industrialized world. Atherosclerosis, the most prevalent of cardiovascular diseases, is the principal cause of heart attack, stroke, and gangrene of the extremities, and thereby the principle cause of death in the United States. Atherosclerosis is a complex disease involving many cell types and molecular factors (described in, for example, Ross, 1993, *Nature* 362: 801-809). The process, in normal circumstances a protective response to insults to the endothelium and smooth muscle cells (SMCs) of the wall of the artery, consists of the formation of fibrofatty and fibrous lesions or plaques, preceded and accompanied by inflammation. The advanced lesions of atherosclerosis may occlude the artery concerned, and result from an excessive inflammatory-fibroproliferative response to numerous different forms of insult. Injury or dysfunction of the vascular endothelium is a common feature of may conditions that predispose an individual to accelerated development of atherosclerotic cardiovascular disease. For example, shear stresses are thought to be responsible for the frequent occurrence of atherosclerotic plaques in regions of the circulatory system where turbulent blood flow occurs, such as branch points and irregular structures.

The first observable event in the formation of an atherosclerotic plaque occurs when blood-borne monocytes adhere to the vascular endothelial layer and transmigrate through to the sub-endothelial space. Adjacent endothelial cells at the same time produce oxidized low density lipoprotein (LDL). These oxidized LDLs are then taken up in large amounts by the monocytes through scavenger receptors expressed on their surfaces. In contrast to the regulated pathway by which native LDL (nLDL) is taken up by nLDL specific receptors, the scavenger pathway of uptake is not regulated by the monocytes.

These lipid-filled monocytes are called foam cells, and are the major constituent of the fatty streak. Interactions between foam cells and the endothelial and SMCs which surround them lead to a state of chronic local inflammation which can eventually lead to smooth muscle cell proliferation and migration, and the formation of a fibrous plaque.

Such plaques occlude the blood vessel concerned and, thus, restrict the flow of blood, resulting in ischemia. Ischemia is a condition characterized by a lack of oxygen supply in tissues of organs due to inadequate perfusion. Such inadequate perfusion can have a number of natural causes, including atherosclerotic or restenotic lesions, anemia, or stroke. Many medical interventions, such as the interruption of the flow of blood during bypass surgery, for example, also lead to ischemia. In addition to sometimes being caused by diseased cardiovascular tissue, ischemia may sometimes affect cardiovascular tissue, such as in ischemic heart disease. Ischemia may occur in any organ, however, that is suffering a lack of oxygen supply.

The most common cause of ischemia in the heart is atherosclerotic disease of epicardial coronary arteries. By reducing the lumen of these vessels, atherosclerosis causes an absolute decrease in myocardial perfusion in the basal state or limits appropriate increases in perfusion when the demand for flow is augmented. Coronary blood flow can also be limited by arterial thrombi, spasm, and, rarely, coronary emboli, as well as by ostial narrowing due to luetic aortitis. Congenital abnormalities, such as anomalous origin of the left anterior descending coronary artery from the pulmonary artery, may cause myocardial ischemia and infarction in infancy, but this cause is very rare in adults.

Myocardial ischemia can also occur if myocardial oxygen demands are abnormally increased, as in severe ventricular hypertrophy due to hypertension or aortic stenosis. The latter can be present with angina that is indistinguishable from that caused by coronary atherosclerosis. A reduction in the oxygen-carrying capacity of the blood, as in extremely severe anemia or in the presence of carboxy-hemoglobin, is a rare cause of myocardial ischemia. Not infrequently, two or more causes of ischemia will coexist, such as an increase in oxygen demand due to left ventricular hypertrophy and a reduction in oxygen supply secondary to coronary atherosclerosis.

The principal surgical approaches to the treatment of ischemic atherosclerosis are bypass grafting, endarterectomy, and percutaneous translumenal angioplasty (PCTA). The failure rate after these approaches due to restenosis, in which the occlusions recur and often become even worse, is extraordinarily high (30-50%). It appears that much of the restenosis is due to further inflammation, smooth muscle accumulation, and thrombosis. Additional therapeutic approaches to cardiovascular disease have included treatments that encouraged angiogenesis in such conditions as ischemic heart and limb disease.

Angiogenesis is a fundamental process by which new blood vessels are formed, as reviewed, for example, by Folkman and Shing, J. Biol. Chem. 267 (16), 10931-10934 (1992). Capillary blood vessels consist of endothelial cells and pericytes. These two cell types carry all of the genetic information to form tubes, branches and whole capillary networks. Specific angiogenic molecules and growth factors can initiate this process. Specific inhibitory molecules can stop it. These molecules with opposing function appear to be continuously acting in concert to maintain a stable microvasculature in which endothelial cell turnover is thousands of days. However, the same endothelial cells can undergo rapid proliferation, *e.g.* in less than five days, during bursts of angiogenesis (for example, during wound healing).

Key components of the angiogenic process are the degradation of the basement membrane, the migration and proliferation of capillary endothelial cell (EC) and the formation of three dimensional capillary tubes. The normal vascular turnover is rather low: the doubling time for capillary endothelium is from 50-20,000 days, but it is 2-13 days for tumor capillary endothelium. The current understanding of the sequence of events leading to angiogenesis is that a cytokine capable of stimulating endothelial cell

proliferation, such as fibroblast growth factor (FGF), causes release of collagenase or plasminogen activator which, in turn, degrade the basement membrane of the parent venule to facilitate the migration of the endothelial cells. These capillary cells, having sprouted from the parent vessel, proliferate in response to growth factors and angiogenic agents in the surrounding environment to form lumen and eventually new blood vessels.

The development of a vascular blood supply is essential in reproduction, development and wound repair (Folkman, *et al.*, Science 43, 1490-1493 (1989)). Under these conditions, angiogenesis is highly regulated, so that it is turned on only as necessary, usually for brief periods of days, then completely inhibited. However, a number of serious diseases are also dominated by persistent unregulated angiogenesis and/or abnormal neovascularization including solid tumor growth and metastasis, psoriasis, endometriosis, Grave's disease, ischemic disease (*e.g.*, atherosclerosis), and chronic inflammatory diseases (*e.g.*, rheumatoid arthritis), and some types of eye disorders, (reviewed by Auerbach, *et al.*, J. Microvasc. Res. 29, 401-411 (1985); Folkman, Advances in Cancer Research, eds. Klein and Weinhouse, pp. 175-203 (Academic Press, New York 1985); Patz, Am. J. Opthalmol. 94, 715-743 (1982); and Folkman, *et al.*, Science 221, 719-725 (1983)). For example, there are a number of eye diseases, many of which lead to blindness, in which ocular neovascularization occurs in response to the diseased state. These ocular disorders include diabetic retinopathy, macular degeneration, neovascular glaucoma, inflammatory diseases and ocular tumors (*e.g.*, retinoblastoma). There are a number of other eye diseases which are also associated with neovascularization, including retrolental fibroplasia, uveitis, eye diseases associated with choroidal neovascularization and eye diseases which are associated with iris neovascularization.

Vascular tone refers to the degree of constriction experienced by a blood vessel relative to its maximal dilated state. All vessels under basal conditions exhibit some degree of smooth muscle contraction that determines the diameter, and hence tone, of the vessel. Basal vascular tone differs among organs wherein organs with a large vasodilatory capacity have high vascular tone (*e.g.*, myocardium, skeletal muscle, skin), and organs with low vasodilatory capacity have low vascular tone (*e.g.*, cerebral and renal circulatory systems).

Vascular tone is determined by many different competing vasoconstrictor and vasodilator influences acting upon the blood vessel. These influences can be separated into extrinsic factors that originate from outside of the organ or tissue where the blood vessel is located, and intrinsic factors that originate from the vessel itself or the surrounding tissue. Extrinsic factors primarily serve the function of regulating arterial blood pressure, while intrinsic mechanisms are concerned with local blood flow regulation within an organ. Vascular tone at any given instant is determined by the balance of competing vasoconstrictor and vasodilator influences.

Summary of the Invention

The present invention is based, at least in part, on the discovery of novel transient receptor potential (TRP) family members, referred to herein as TRP-like calcium channel or TLCC nucleic acid and protein molecules. The TLCC molecules of the present invention are useful as targets for developing modulating agents to regulate a variety of cellular processes, including contractility of cells, such as stellate cells, membrane excitability, neurite outgrowth and synaptogenesis, signal transduction, cell proliferation, growth, differentiation, and migration, and nociception. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding TLCC proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of TLCC-encoding nucleic acids.

The present invention is also based, at least in part, on the discovery that the TLCC gene is up-regulated in stellate cells (the main effectors of liver fibrosis) as compared to its expression in normal hepatic cells, and, thus, may be associated with a hepatic disorder. Accordingly, the present invention also provides methods and compositions for the diagnosis and treatment of a hepatic disorder, including but not limited to, liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis.

The present invention is further based, at least in part, on the discovery that TLCC expression is regulated by two stimuli relevant to atherosclerosis and angiogenesis, IL-1β and shear stress. Specifically, the present invention demonstrates

that the TLCC gene is expressed in human blood vessels and endothelial cells (Example 4), that the TLCC gene expression in endothelial cells is down-regulated when endothelial cells are treated with IL-1β (Example 4), and that the TLCC gene is upregulated in endothelial cells treated under conditions of laminar shear stress (LSS). Accordingly, the present invention also provides methods and compositions for the diagnosis and treatment of cardiovascular disease, including but not limited to, atherosclerosis, ischemia/reperfusion injury, hypertension, restenosis, arterial inflammation, and endothelial cell disorders, such as disorders associated with aberrant endothelial cell growth, angiogenesis and/or vascularization.

In one embodiment, a TLCC nucleic acid molecule of the invention is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1 or 3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1 or 3, or a complement thereof. In another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 1-137 of SEQ ID NO:1. In yet another embodiment, the nucleic acid molecule includes SEQ ID NO:3 and nucleotides 3529-3900 of SEQ ID NO:1. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1 or 3. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 2393 nucleotides (*e.g.*, 2393 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO: 1or 3, or a complement thereof.

In another embodiment, a TLCC nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a preferred embodiment, a TLCC nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the amino acid sequence of SEQ ID NO:2 or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human TLCC. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, the nucleic acid molecule is at least 2393 nucleotides in length. In a further preferred embodiment, the nucleic acid molecule is at least 2393 nucleotides in length and encodes a protein having a TLCC activity (as described herein).

Another embodiment of the invention features nucleic acid molecules, preferably TLCC nucleic acid molecules, which specifically detect TLCC nucleic acid molecules relative to nucleic acid molecules encoding non-TLCC proteins. For example, in one embodiment, such a nucleic acid molecule is at least 1767, 1767-1800, 1800-1900, 1900-2000, 2000-2100, 2100-2200, 2200-2300, 2300-2393, 2393-2400, 2400-2500, 2500-2600, 2600-2700, 2700-2800, 2800-2900, 2900-3000 or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number_____, or a complement thereof.

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 3 under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a TLCC nucleic acid molecule, *e.g.*, the coding strand of a TLCC nucleic acid molecule.

Another aspect of the invention provides a vector comprising a TLCC nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. In yet another embodiment, the invention provides a host cell containing a nucleic acid molecule of the invention. The invention also provides a method for

producing a protein, preferably a TLCC protein, by culturing in a suitable medium a host cell, *e.g.*, a mammalian host cell such as a liver cell, of the invention containing a recombinant expression vector, such that the protein is produced.

Another aspect of this invention features isolated or recombinant TLCC proteins and polypeptides. In one embodiment, an isolated TLCC protein includes at least one transmembrane domain. In another embodiment, an isolated TLCC protein includes at least one N-glycosylation site. In yet another embodiment, an isolated TLCC protein includes at least one transmembrane calcium channel domain. In yet another embodiment, an isolated TLCC protein includes at least one transmembrane domain and one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain. In yet another embodiment, an isolated TLCC protein includes at least one transmembrane domain, at least one N-glycosylation site, and a transmembrane calcium channel domain.

In a preferred embodiment, a TLCC protein includes at least one transmembrane domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a further preferred embodiment, a TLCC protein includes a transmembrane calcium channel domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In a further preferred embodiment, a TLCC protein includes at least one transmembrane domain and one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

In another preferred embodiment, a TLCC protein includes at least one transmembrane domain and has a TLCC activity (as described herein).

In yet another preferred embodiment, a TLCC protein includes at least one transmembrane domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3. In a further embodiment, a TLCC protein includes a transmembrane calcium channel domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3. In another embodiment, a TLCC protein includes at least one transmembrane domain and one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 or 3.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 15, 50, 100, 150, 200, 250, 300, 315, 316, 325, 350, 375, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1126 or more amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number_____. In another embodiment, a TLCC protein has the amino acid sequence of SEQ ID NO:2.

In another embodiment, the invention features a TLCC protein which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof. This invention further features a TLCC protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

The proteins of the present invention or portions thereof, *e.g.*, biologically active portions thereof, can be operatively linked to a non-TLCC polypeptide (*e.g.*, heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind proteins of the invention, preferably TLCC proteins. In addition, the TLCC proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting the presence of a TLCC nucleic acid molecule, protein, or polypeptide in a biological sample by contacting the biological sample with an agent capable of detecting a TLCC nucleic acid molecule, protein, or polypeptide such that the presence of a TLCC nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the presence of TLCC activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of TLCC activity such that the presence of TLCC activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating TLCC activity comprising contacting a cell capable of expressing TLCC with an agent that modulates TLCC activity such that TLCC activity in the cell is modulated. In one embodiment, the agent inhibits TLCC activity. In another embodiment, the agent stimulates TLCC activity. In one embodiment, the agent is an antibody that specifically binds to a TLCC protein. In another embodiment, the agent modulates expression of TLCC by modulating transcription of a TLCC gene or translation of a TLCC mRNA. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a TLCC mRNA or a TLCC gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant or unwanted TLCC protein or nucleic acid expression or activity by administering an agent which is a TLCC modulator to the subject. In one embodiment, the TLCC modulator is a TLCC protein. In another embodiment the TLCC modulator is a TLCC nucleic acid molecule. In yet another embodiment, the TLCC modulator is a peptide, peptidomimetic, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant or

unwanted TLCC activity is a hepatic disorder (such as, liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis).

In another preferred embodiment, the disorder characterized by aberrant or unwanted TLCC activity is a cardiovascular disorder or an endothelial cell disorder.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TLCC protein; (ii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TLCC protein, wherein a wild-type form of the gene encodes a protein with a TLCC activity.

In another aspect the invention provides methods for identifying a compound that binds to or modulates the activity of a TLCC protein, by providing an indicator composition comprising a TLCC protein having TLCC activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on TLCC activity in the indicator composition to identify a compound that modulates the activity of a TLCC protein.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figures 1A-1D depict the cDNA sequence and predicted amino acid sequence of human TLCC. The nucleotide sequence corresponds to nucleic acids 1 to 3900 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 1130 of SEQ ID NO: 2. The coding region without the 3' untranslated region of the human TLCC gene is shown in SEQ ID NO:3.

Figure 2 depicts a structural, hydrophobicity, and antigenicity analysis of the human TLCC protein.

Figures 3A-3E depict an alignment of the nucleotide sequence of human TLCC with the nucleotide sequence of human PS112 consensus DNA fragment from gene specific clones (Accession Number V26656), using the CLUSTAL W (1.74) multiple sequence alignment program.

Figures 4A-4D depict an alignment of the translated human TLCC cDNA sequence with the amino acid sequences of 'similar to *C. elegans* hypothetical protein CET01H8.1, CEC05C12.3, CEF54D1.5, similar to trp and trp-like proteins' (Accession No. BAA34700), of *Homo sapiens* melastatin I (Accession No. AAC80000), and of 'similarity with Drosophila transient-receptor-potential protein (Swiss Prot accession number P19334); cDNA EST EMBL:D27562 comes from this gene; cDNA EST yk219f12.5 comes from this gene [*Caenorhabditis elegans*] (Accession No. CAB05572) from *Homo sapiens*, using the CLUSTAL W (1.74) multiple sequence alignment program.

Figures 5A-5C depict an alignment of the translated human TLCC cDNA sequence with the amino acid sequences of human PS112 protein sequence from gene-specific clones (Accession Number W54425), with prostate tumour specific gene clone J1-17 protein (Accession Number W69384), with 'amino acid encoded by prostate tumour clone J1-17' (Accession Number W71868), and with 'prostate tumour derived antigen #4' (Accession Number Y00931), using the CLUSTAL W (1.74) multiple sequence alignment program.

Figure 6 depicts the results of a search which was performed against the MEMSAT database and which resulted in the identification of six "transmembrane domains" in the human TLCC protein.

Figure 7 depicts the results of a search which was performed against the Prosite database and which resulted in the identification of three N-glycosylation sites in the human TLCC protein.

Figures 8A-8B depict the results of a search which was performed against the ProDom database and which resulted in the identification of four regions of similarity to human melastatin and a "transmembrane calcium channel domain" in the human TLCC protein.

Figures 9A-9C depict an alignment of the human TLCC amino acid sequence with the amino acid sequence of human melastatin (Accession Number AAC80000), using the GAP program in the GCG software package (Blosum 62 matrix), a gap weight of 12, and a length weight of 4.

Figure 10 depicts the results of RT-PCR analysis of human TLCC expression in various vessels and cells derived therefrom.

Figure 11 depicts human TLCC expression in endothelial cells during laminar shear stress.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as " TRP-like calcium channel" or "TLCC" nucleic acid and protein molecules, which are novel members of the calcium channel family. These novel molecules are capable of, for example, modulating a calcium channel mediated activity in a cell, *e.g.*, a neuronal, muscle (*e.g.*, cardiac muscle), or liver cell. The present invention is further based, at least in part, on the discovery that TLCC genes are up-regulated in stellate cells (the main effectors of liver fibrosis) as compared to their expression in normal hepatic cells, and, thus, may be associated with a hepatic disorder. Accordingly, the present invention further provides methods and compositions for the diagnosis and treatment of a hepatic disorder, including but not limited to, liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis.

As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, *e.g.*, a neuronal or muscle cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, *e.g.*, neuronal cells, and may form heteromultimeric structures (*e.g.*, composed of more than one type of subunit). Calcium channels may also be found in nonexcitable cells (*e.g.*, adipose cells or liver cells), where they may play a role in, *e.g.*, signal transduction. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila *et al.* (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M.W. *et al.* (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference. As the TLCC molecules of the present invention may modulate calcium channel mediated

activities, they may be useful for developing novel diagnostic and therapeutic agents for calcium channel associated disorders.

As used herein, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of calcium channel mediated activity. Calcium channel associated disorders include cardiovascular disease and hepatic disorders.

As used herein, "cardiovascular disease" or a "cardiovascular disorder" includes a disease or disorder which affects the cardiovascular system, *e.g.*, the heart or the blood vessels. A cardiovascular disorder includes disorders such as arteriosclerosis, ischemia reperfusion injury, restenosis, arterial inflammation, vascular wall remodeling, ventricular remodeling, rapid ventricular pacing, coronary microembolism, tachycardia, bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, atrial fibrillation, long-QT syndrome, congestive heart failure, sinus node dysfunction, angina, heart failure, hypertension, atrial fibrillation, atrial flutter, dilated cardiomyopathy, idiopathic cardiomyopathy, myocardial infarction, coronary artery disease, coronary artery spasm, ischemic disease, arrhythmia, and cardiovascular developmental disorders (*e.g.*, arteriovenous malformations, arteriovenous fistulae, Raynaud's syndrome, neurogenic thoracic outlet syndrome, causalgia/reflex sympathetic dystrophy, hemangioma, aneurysm, cavernous angioma, aortic valve stenosis, atrial septal defects, atrioventricular canal, coarctation of the aorta, ebsteins anomaly, hypoplastic left heart syndrome, interruption of the aortic arch, mitral valve prolapse, ductus arteriosus, patent foramen ovale, partial anomalous pulmonary venous return, pulmonary atresia with ventricular septal defect, pulmonary atresia without ventricular septal defect, persistance of the fetal circulation, pulmonary valve stenosis, single ventricle, total anomalous pulmonary venous return, transposition of the great vessels, tricuspid atresia, truncus arteriosus, ventricular septal defects). A cardiovasular disease or disorder also includes an endothelial cell and/or smooth muscle cell disorder. As used herein, an "endothelial cell disorder" and/or a "smooth muscle cell disorder" includes a disorder characterized by aberrant, unregulated, or unwanted endothelial cell activity, *e.g.*, vascular tone, vasodilation, vasoconstriction, proliferation, migration, angiogenesis, or vascularization; or aberrant expression of cell surface adhesion molecules or genes associated with angiogenesis, *e.g.*, TIE-2, FLT and FLK.

Endothelial cell disorders include tumorigenesis, tumor metastasis, psoriasis, diabetic retinopathy, endometriosis, Grave's disease, ischemic disease (*e.g.*, atherosclerosis), chronic inflammatory diseases (*e.g.*, rheumatoid arthritis), arterial hypertension, pulmonary hypertension, primary pulmonary hypertension (PPH), Raynaud's phenomenon (RP), migraine headache, chronic heart failure, erythromelalgia, familial dysautonomia, hemolytic uremic syndrome, preeclampsia, reperfusion injury, postangioplasty enothelial regeneration, degeneration of venous bypass grafts, angina, pure spastic angina, diabetes, reflex sympathetic dystrophy syndrome, and vasculitis.

As used herein, a "hepatic disorder" includes a disorder, disease or condition which affects the liver. The term hepatic disorder includes a disorder caused by the over- or under-production of hepatic enzymes, *e.g.*, alanine aminotransferase, aspartate aminotransferase, or γ-glutammyl transferase, in the liver. For example, a hepatic disorder includes hepatic fibrosis, a hepatic disorder caused by a drug, a hepatic disorder caused by prolonged ethanol uptake, a hepatic injury caused by carbon tetrachloride exposure, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, or autoimmune hepatitis. Hepatic disorders are disclosed at, for example, the American Liver Foundation website (found at the world wide web address: gi.ucsf.edu/alf.html).

A hepatic disorder also includes a hepatic cell disorder. As used herein a "hepatic cell disorder" includes a disorder characterized by aberrant or unwanted hepatic cell activity, *e.g.*, proliferation, migration, angiogenesis, or aberrant expression of cell surface adhesion molecules.

Calcium channel disorders may also include CNS disorders, such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Creutzfeldt-Jakob disease, or AIDS related dementia; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression , schizophrenia, schizoaffective disorder, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; leaning or memory disorders, *e.g.*, amnesia or age-related memory loss,

attention deficit disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

Calcium channel disorders also include pain disorders. Pain disorders include those that affect pain signaling mechanisms. As used herein, the term "pain signaling mechanisms" includes the cellular mechanisms involved in the development and regulation of pain, *e.g.*, pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, *e.g.*, a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter sensory neurons. These sensory neurons transmit the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes. The TLCC molecules of the present invention may be present on these sensory neurons and, thus, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the TLCC molecules by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

Calcium channel disorders also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes. As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. The TLCC molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TLCC molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders

characterized by aberrantly regulated growth, differentiation, or migration. Such disorders include cancer, *e.g.*, carcinoma, sarcoma, or leukemia; tumor angiogenesis and metastasis; skeletal dysplasia; neuronal deficiencies resulting from impaired neural induction and patterning; hepatic disorders; cardiovascular disorders; and hematopoietic and/or myeloproliferative disorders.

As used herein, a "calcium channel mediated activity" includes an activity which involves a calcium channel, *e.g.*, a calcium channel in a neuronal cell, a muscular cell, a vascular cell, or a liver cell, associated with receiving, conducting, and transmitting signals, in, for example, the nervous system. Calcium channel mediated activities include release of neurotransmitters or second messenger molecules (*e.g.*, dopamine or norepinephrine), from cells, *e.g.*, neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (*e.g.*, changes in those action potentials resulting in a morphological or differentiative response in the cell).

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin, *e.g.*, monkey proteins. Members of a family may also have common functional characteristics.

For example, the family of TLCC proteins comprise at least one "transmembrane domain" and preferably six transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred

embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, tyrosines, or tryptophans. Transmembrane domains are described in, for example, Zagotta W.N. *et al,* (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference. Amino acid residues 599-619, 690-712, 784-803, 811-831, 845-862, and 933-957 of the TLCC protein comprise transmembrane domains (see Figures 2 and 6). Accordingly, TLCC proteins having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC are within the scope of the invention.

In another embodiment, a TLCC molecule of the present invention is identified based on the presence of at least one pore domain between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described, for example in Vannier *et al.* (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. *et al.* (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. Amino acid residues 880-900 of the TLCC protein comprise a pore domain (see Figures 2 and 6).

In another embodiment, a TLCC molecule of the present invention is identified based on the presence of at least one N-glycosylation site. As used herein, the term "N-glycosylation site" includes an amino acid sequence of about 4 amino acid residues in length which serves as a glycosylation site. More preferably, an N-glycosylation site has the consensus sequence Asn-Xaa-Ser/Thr (where Xaa may be any amino acid) (SEQ ID NO:4). N-glycosylation sites are described in, for example, Prosite PDOC00001 (found at the world wide web address: expasy.ch/cgi-bin/get-prodoc-entry?PDOC00001), the contents of which are incorporated herein by reference. Amino acid residues 143-146, 205-208, and 907-910 of the TLCC protein comprise N-glycosylation sites (see Figure 7). Accordingly, TLCC proteins having at least one N-glycosylation site are within the scope of the invention.

In another embodiment, a TLCC molecule of the present invention is identified based on the presence of a "transmembrane calcium channel domain" in the protein or corresponding nucleic acid molecule. As used herein, the term "transmembrane calcium channel domain" includes a protein domain having an amino acid sequence of about 40-100 amino acid residues and having a bit score for the alignment of the sequence to the transmembrane calcium channel domain of at about 50-100. Preferably, a transmembrane calcium channel domain includes at least about 60-80, or more preferably about 63 amino acid residues, and has a bit score for the alignment of the sequence to the transmembrane calcium channel domain of at least 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 or higher. The transmembrane calcium channel domain has been assigned ProDom entry 2328. To identify the presence of a transmembrane calcium channel domain in a TLCC protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein is searched against a database of known protein domains (*e.g.*, the ProDom database) using the default parameters (available at the world wide web address: toulouse.inra.fr/prodom.html). A search was performed against the ProDom database resulting in the identification of a transmembrane calcium channel domain in the amino acid sequence of human TLCC (SEQ ID NO: 2) at about residues 783-845 of SEQ ID NO: 2. The results of the search are set forth in Figure 8.

Isolated proteins of the present invention, preferably TLCC proteins, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least 30%, 40%, or 50% homology, preferably 60% homology, more preferably 70%-80%, and even more preferably 90-95% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide

sequences which share at least 30%, 40%, or 50%, preferably 60%, more preferably 70-80%, or 90-95% homology and share a common functional activity are defined herein as sufficiently identical.

As used interchangeably herein, an "TLCC activity", "biological activity of TLCC " or "functional activity of TLCC", refers to an activity exerted by a TLCC protein, polypeptide or nucleic acid molecule on a TLCC responsive cell or tissue, or on a TLCC protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a TLCC activity is a direct activity, such as an association with a TLCC-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC protein binds or interacts in nature, such that TLCC-mediated function is achieved. A TLCC target molecule can be a non-TLCC molecule or a TLCC protein or polypeptide of the present invention. In an exemplary embodiment, a TLCC target molecule is a TLCC ligand, *e.g.*, a calcium channel ligand. Alternatively, a TLCC activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC protein with a TLCC ligand. The biological activities of TLCC are described herein. For example, the TLCC proteins of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, (7) participate in nociception, (8) modulate hepatic disorders, (9) modulate angiogenesis, (10) modulate endothelial cell proliferation, and (11) modulate vascular tone.

Accordingly, another embodiment of the invention features isolated TLCC proteins and polypeptides having a TLCC activity. Preferred proteins are TLCC proteins having at least one transmembrane domain, and, preferably, a TLCC activity. Other preferred proteins are TLCC proteins having an N-glycosylation site and, preferably, a TLCC activity. Yet other preferred proteins are TLCC proteins having at least one transmembrane calcium channel domain and, preferably, a TLCC activity. Yet other preferred proteins are TLCC proteins having at least one transmembrane domain, at least one N-glycosylation site, and a transmembrane calcium channel domain and, preferably, a TLCC activity.

Additional preferred proteins have at least one transmembrane domain, and one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

The nucleotide sequence of the isolated human TLCC cDNA and the predicted amino acid sequence of the human TLCC polypeptide are shown in Figure 1 and in SEQ ID NOs:1 and 2, respectively. A plasmid containing the nucleotide sequence encoding human TLCC was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human TLCC gene, which is approximately 3900 nucleotides in length, encodes a protein having a molecular weight of approximately 128 kD and which is approximately 1130 amino acid residues in length.

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode TLCC proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify TLCC-encoding nucleic acid molecules (*e.g.*, TLCC mRNA) and fragments for use as PCR primers for the amplification or mutation of TLCC nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TLCC nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, TLCC nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TLCC nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human TLCC cDNA. This cDNA comprises sequences encoding the human TLCC protein (*i.e.*, "the coding region", from nucleotides 138-3528), as well as 5' untranslated sequences (nucleotides 1-137) and 3' untranslated sequences (nucleotides 3529-3900). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 138-3528, corresponding to SEQ ID NO:3).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ , or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the entire

length of the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a TLCC protein, *e.g.*, a biologically active portion of a TLCC protein. The nucleotide sequence determined from the cloning of the TLCC gene allows for the generation of probes and primers designed for use in identifying and/or cloning other TLCC family members, as well as TLCC homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, or 100 or more consecutive nucleotides of a sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is greater than 1767, 1767-1800, 1800-1900, 1900-2000, 2000-2100, 2100-2200, 2200-2300, 2300-2393, 2393-2400, 2400-2500, 2500-2600, 2600-2700, 2700-2800, 2800-2900, 2900-3000 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Probes based on the TLCC nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or

tissue which misexpress a TLCC protein, such as by measuring a level of a TLCC-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting TLCC mRNA levels or determining whether a genomic TLCC gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a TLCC protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a TLCC biological activity (the biological activities of the TLCC proteins are described herein), expressing the encoded portion of the TLCC protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the TLCC protein.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same TLCC proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2.

In addition to the TLCC nucleotide sequences shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the TLCC proteins may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the TLCC genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a TLCC protein, preferably a mammalian TLCC protein, and can further include non-coding regulatory sequences, and introns.

Allelic variants of human TLCC include both functional and non-functional TLCC proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC protein that maintain the ability to bind a TLCC ligand or

substrate and/or modulate membrane excitability or signal transduction. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC protein that do not have the ability to form functional calcium channels or to modulate membrane excitability. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human TLCC proteins. Orthologues of the human TLCC protein are proteins that are isolated from non- non-human organisms and possess the same TLCC ligand binding and/or modulation of membrane excitation mechanisms of the human TLCC protein. Orthologues of the human TLCC protein can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2.

Moreover, nucleic acid molecules encoding other TLCC family members and, thus, which have a nucleotide sequence which differs from the TLCC sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another TLCC cDNA can be identified based on the nucleotide sequence of human TLCC. Moreover, nucleic acid molecules encoding TLCC proteins from different species, and which, thus, have a nucleotide sequence which differs from the TLCC sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse TLCC cDNA can be identified based on the nucleotide sequence of a human TLCC.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC cDNAs of the invention can be isolated based on their homology to the TLCC nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural

allelic variants and homologues of the TLCC cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the TLCC gene.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic acid is at least 1767, 1767-1800, 1800-1900, 1900-2000, 2000-2100, 2100-2200, 2200-2300, 2300-2393, 2393-2400, 2400-2500, 2500-2600, 2600-2700, 2700-2800, 2800-2900, 2900-3000 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or alternatively hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or alternatively hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE

is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 each minutes after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

In addition to naturally-occurring allelic variants of the TLCC sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded TLCC proteins, without altering the functional ability of the TLCC proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of TLCC (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is

required for biological activity. For example, amino acid residues that are conserved among the TLCC proteins of the present invention, *e.g.*, those present in a transmembrane domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TLCC proteins of the present invention and other members of the TLCC family are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TLCC proteins that contain changes in amino acid residues that are not essential for activity. Such TLCC proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a TLCC protein identical to the protein of SEQ ID NO:2, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino

acid residue in a TLCC protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a TLCC coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TLCC biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant TLCC protein can be assayed for the ability to (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, (7) participate in nociception, (8) modulate hepatic disorders, (9) modulate angiogenesis, (10) modulate endothelial cell proliferation, and (11) modulate vascular tone.

In addition to the nucleic acid molecules encoding TLCC proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire TLCC coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TLCC. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human TLCC corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TLCC. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding TLCC disclosed herein (*e.g.*, SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TLCC mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of TLCC mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TLCC mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TLCC protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TLCC mRNA transcripts to thereby inhibit translation of TLCC mRNA. A ribozyme having specificity for a TLCC-encoding nucleic acid can be designed based upon the nucleotide sequence of a TLCC cDNA disclosed herein (*i.e.*,

SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TLCC-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, TLCC mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, TLCC gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TLCC (*e.g.*, the TLCC promoter and/or enhancers; *e.g.*, nucleotides 1-137 of SEQ ID NO:1) to form triple helical structures that prevent transcription of the TLCC gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the TLCC nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of TLCC nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of TLCC nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-

directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of TLCC can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TLCC nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the

oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

Alternatively, the expression characteristics of an endogenous TLCC gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous TLCC gene. For example, an endogenous TLCC gene which is normally "transcriptionally silent", *i.e.*, a TLCC gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous TLCC gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous TLCC gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

II. Isolated TLCC Proteins and Anti-TLCC Antibodies

One aspect of the invention pertains to isolated TLCC proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-TLCC antibodies. In one embodiment, native TLCC proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TLCC proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a TLCC protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TLCC protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language

"substantially free of cellular material" includes preparations of TLCC protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of TLCC protein having less than about 30% (by dry weight) of non-TLCC protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TLCC protein, still more preferably less than about 10% of non-TLCC protein, and most preferably less than about 5% non-TLCC protein. When the TLCC protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC protein having less than about 30% (by dry weight) of chemical precursors or non-TLCC chemicals, more preferably less than about 20% chemical precursors or non-TLCC chemicals, still more preferably less than about 10% chemical precursors or non-TLCC chemicals, and most preferably less than about 5% chemical precursors or non-TLCC chemicals.

As used herein, a "biologically active portion" of a TLCC protein includes a fragment of a TLCC protein which participates in an interaction between a TLCC molecule and a non-TLCC molecule. Biologically active portions of a TLCC protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the TLCC protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length TLCC proteins, and exhibit at least one activity of a TLCC protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the TLCC protein, *e.g.*, modulating membrane excitation mechanisms. A biologically active portion of a TLCC protein can be a polypeptide which is, for example, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 316, 325, 350, 375, 400, 425, 450, 274, 500, 525, 550, 575, 600, 625, 650, 675, or 700 or more amino acids in length. Biologically active

portions of a TLCC protein can be used as targets for developing agents which modulate a TLCC mediated activity, *e.g.*, a membrane excitation mechanism.

In one embodiment, a biologically active portion of a TLCC protein comprises at least one transmembrane domain. It is to be understood that a preferred biologically active portion of a TLCC protein of the present invention comprises at least one transmembrane domain and may additionally contain one or more of the following domains: at least one N-glycosylation site, and a transmembrane calcium channel domain. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TLCC protein.

In a preferred embodiment, the TLCC protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the TLCC protein is substantially identical to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the TLCC protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the TLCC amino acid sequence of SEQ ID NO:2 having 1130 amino acid residues, at least 50, preferably at least 100, 200, 300, 400, 500, more preferably at least 600, 700, 800, even more preferably at least 900, and even more preferably at least 1000, 1050, 1100 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid

residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at the world wide web address: gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at the world wide web address: gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.*, 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0 or 2.0U) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TLCC nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 100, wordlength = 3, and a Blosum62 matrix to obtain amino acid sequences homologous to TLCC protein molecules of the invention. To obtain gapped alignments for comparison purposes,

Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See the world wide web address: ncbi.nlm.nih.gov.

The invention also provides TLCC chimeric or fusion proteins. As used herein, a TLCC "chimeric protein" or "fusion protein" comprises a TLCC polypeptide operatively linked to a non-TLCC polypeptide. An "TLCC polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TLCC, whereas a "non-TLCC polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the TLCC protein, *e.g.*, a protein which is different from the TLCC protein and which is derived from the same or a different organism. Within a TLCC fusion protein the TLCC polypeptide can correspond to all or a portion of a TLCC protein. In a preferred embodiment, a TLCC fusion protein comprises at least one biologically active portion of a TLCC protein. In another preferred embodiment, a TLCC fusion protein comprises at least two biologically active portions of a TLCC protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the TLCC polypeptide and the non-TLCC polypeptide are fused in-frame to each other. The non-TLCC polypeptide can be fused to the N-terminus or C-terminus of the TLCC polypeptide.

For example, in one embodiment, the fusion protein is a GST-TLCC fusion protein in which the TLCC sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant TLCC.

In another embodiment, the fusion protein is a TLCC protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of TLCC can be increased through use of a heterologous signal sequence.

The TLCC fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TLCC fusion proteins can be used to affect the bioavailability of a TLCC substrate. Use of TLCC fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TLCC protein;

(ii) mis-regulation of the TLCC gene; and (iii) aberrant post-translational modification of a TLCC protein.

Moreover, the TLCC-fusion proteins of the invention can be used as immunogens to produce anti-TLCC antibodies in a subject, to purify TLCC ligands and in screening assays to identify molecules which inhibit the interaction of TLCC with a TLCC substrate.

Preferably, a TLCC chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A TLCC-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TLCC protein.

The present invention also pertains to variants of the TLCC proteins which function as either TLCC agonists (mimetics) or as TLCC antagonists. Variants of the TLCC proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a TLCC protein. An agonist of the TLCC proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a TLCC protein. An antagonist of a TLCC protein can inhibit one or more of the activities of the naturally occurring form of the TLCC protein by, for example, competitively modulating a TLCC-mediated activity of a TLCC protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological

activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the TLCC protein.

In one embodiment, variants of a TLCC protein which function as either TLCC agonists (mimetics) or as TLCC antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a TLCC protein for TLCC protein agonist or antagonist activity. In one embodiment, a variegated library of TLCC variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TLCC variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TLCC sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of TLCC sequences therein. There are a variety of methods which can be used to produce libraries of potential TLCC variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TLCC sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of a TLCC protein coding sequence can be used to generate a variegated population of TLCC fragments for screening and subsequent selection of variants of a TLCC protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TLCC coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the TLCC protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TLCC proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TLCC variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated TLCC library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, an endothelial cell line, which ordinarily responds to TLCC in a particular TLCC substrate-dependent manner. The transfected cells are then contacted with TLCC and the effect of expression of the mutant on signaling by the TLCC substrate can be detected, *e.g.*, by monitoring intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the TLCC substrate, and the individual clones further characterized.

An isolated TLCC protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind TLCC using standard techniques for polyclonal and monoclonal antibody preparation. A full-length TLCC protein can be used or, alternatively, the invention provides antigenic peptide fragments of TLCC for use as immunogens. The antigenic peptide of TLCC comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of TLCC such that an antibody raised against the peptide forms a specific immune complex with TLCC. Preferably, the antigenic peptide comprises at least 10

amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of TLCC that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 2).

A TLCC immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed TLCC protein or a chemically synthesized TLCC polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic TLCC preparation induces a polyclonal anti-TLCC antibody response.

Accordingly, another aspect of the invention pertains to anti-TLCC antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TLCC. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind TLCC. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TLCC. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TLCC protein with which it immunoreacts.

Polyclonal anti-TLCC antibodies can be prepared as described above by immunizing a suitable subject with a TLCC immunogen. The anti-TLCC antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized TLCC. If desired, the antibody molecules directed against TLCC can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as

protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-TLCC antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem* .255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a TLCC immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds TLCC.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TLCC monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These

myeloma lines are available from ATCC Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TLCC, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TLCC antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with TLCC to thereby isolate immunoglobulin library members that bind TLCC. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

Additionally, recombinant anti-TLCC antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of

the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-TLCC antibody (*e.g.*, monoclonal antibody) can be used to isolate TLCC by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TLCC antibody can facilitate the purification of natural TLCC from cells and of recombinantly produced TLCC expressed in host cells. Moreover, an anti-TLCC antibody can be used to detect TLCC protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TLCC protein. Anti-TLCC antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include

luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a TLCC protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory

sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, TLCC proteins, mutant forms of TLCC proteins, fusion proteins, and the like).

The recombinant expression vectors of the invention can be designed for expression of TLCC proteins in prokaryotic or eukaryotic cells. For example, TLCC proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in TLCC activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for TLCC proteins, for example. In a preferred embodiment, a TLCC fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the TLCC expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, TLCC proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No.

4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to TLCC mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a TLCC nucleic acid molecule of the invention is introduced, *e.g.*, a TLCC nucleic acid molecule within a recombinant expression vector or a TLCC nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a TLCC protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO), COS cells, or human

umbilical vein endothelial cells (HUVEC)). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a TLCC protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a TLCC protein. Accordingly, the invention further provides methods for producing a TLCC protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a TLCC protein has been introduced) in a suitable medium such that a TLCC protein is produced. In another embodiment, the method further comprises isolating a TLCC protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TLCC-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TLCC sequences have been introduced into their genome or homologous recombinant animals in which endogenous TLCC sequences have been altered. Such animals are useful for studying the function and/or activity of a TLCC and for identifying and/or evaluating modulators of TLCC activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous TLCC gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a TLCC-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TLCC cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TLCC gene, such as a mouse or rat TLCC gene, can be used as a transgene. Alternatively, a TLCC gene homologue, such as another TLCC family member, can be isolated based on hybridization to the TLCC cDNA sequences of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific

regulatory sequence(s) can be operably linked to a TLCC transgene to direct expression of a TLCC protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a TLCC transgene in its genome and/or expression of TLCC mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TLCC protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TLCC gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the TLCC gene. The TLCC gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a non-human homologue of a human TLCC gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1). For example, a mouse TLCC gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous TLCC gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous TLCC gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous TLCC gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous TLCC protein). In the homologous recombination nucleic acid molecule, the altered portion of the TLCC gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the TLCC gene to allow for homologous recombination to occur between the exogenous TLCC gene carried by the homologous recombination nucleic

acid molecule and an endogenous TLCC gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking TLCC nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced TLCC gene has homologously recombined with the endogenous TLCC gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a

transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The TLCC nucleic acid molecules, fragments of TLCC proteins, and anti-TLCC antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils,

polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a TLCC protein or an anti-TLCC antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions

are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e*,. including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is

furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet

of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.*

(eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, a TLCC protein of the invention has one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials,

(4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, (7) participate in nociception, (8) modulate hepatic disorders, (9) modulate angiogenesis, (10) modulate endothelial cell proliferation, and (11) modulate vascular tone. The isolated nucleic acid molecules of the invention can be used, for example, to express TLCC protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect TLCC mRNA (*e.g.*, in a biological sample) or a genetic alteration in a TLCC gene, and to modulate TLCC activity, as described further below. The TLCC proteins can be used to treat disorders characterized by insufficient or excessive production of a TLCC substrate or production of TLCC inhibitors. In addition, the TLCC proteins can be used to screen for naturally occurring TLCC substrates, to screen for drugs or compounds which modulate TLCC activity, as well as to treat disorders characterized by insufficient or excessive production of TLCC protein or production of TLCC protein forms which have decreased, aberrant or unwanted activity compared to TLCC wild type protein (*e.g.*, hepatic disorders such as liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis, CNS disorders such as neurodegenerative disorders, pain disorders, or disorders of cellular growth, differentiation, or migration). Moreover, the anti-TLCC antibodies of the invention can be used to detect and isolate TLCC proteins, to regulate the bioavailability of TLCC proteins, and modulate TLCC activity.

A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to TLCC proteins, have a stimulatory or inhibitory effect on, for example, TLCC expression or TLCC activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of TLCC substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TLCC protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a TLCC protein

or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a TLCC protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TLCC activity is determined. Determining the ability of the test compound to modulate TLCC activity can be accomplished by monitoring, for example, hepatic cell proliferation, contractility, production of extracellular matrix (ECM) components, intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, the activity of a TLCC-regulated transcription factor, or gene expression of, for example, cell surface adhesion molecules or genes associated with angiogenesis. The cell, for

example, can be of a mammalian origin, *e.g.*, a neuronal cell, an endothelial cell, or a liver cell, *e.g.*, HEPG2 cells.

Assays to monitor hepatic TLCC activities include assays based on the nuclear incorporation of 5-bromodeoxyuridine and other colorimetric assays that quantitate cell proliferation; fura-2 and morphometric assays that measure contractility; and the use of antibodies against ECM components (*e.g.*, using ELISA assays) to detect the production of ECM components. These assays are known in the art and are described in, for example, Casini *et al.* (1993) *Gastroenterology* 105:245-253; Gorbig *et al.* (1999) *Hepatology* 30:501-509; Ito *et al.* (2000) *Oncology* 58:261-270; You *et al.* (2000) *Chung Hua Kan Tsang Ping Tsa Chih* 20:78-80; Iwamoto *et al.* (2000) *J. Hepatol.* 32:762-770; Bataller (2000) *Gastroenterology* 118:1149-1156.

The ability of the test compound to modulate TLCC binding to a substrate or to bind to TLCC can also be determined. Determining the ability of the test compound to modulate TLCC binding to a substrate can be accomplished, for example, by coupling the TLCC substrate with a radioisotope or enzymatic label such that binding of the TLCC substrate to TLCC can be determined by detecting the labeled TLCC substrate in a complex. Alternatively, TLCC could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate TLCC binding to a TLCC substrate in a complex. Determining the ability of the test compound to bind TLCC can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TLCC can be determined by detecting the labeled TLCC compound in a complex. For example, compounds (*e.g.*, TLCC substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a TLCC substrate) to interact with TLCC without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with TLCC without the labeling of either the compound or the TLCC. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a

"microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TLCC.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a TLCC target molecule (*e.g.*, a TLCC substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC target molecule. Determining the ability of the test compound to modulate the activity of a TLCC target molecule can be accomplished, for example, by determining the ability of the TLCC protein to bind to or interact with the TLCC target molecule.

Determining the ability of the TLCC protein, or a biologically active fragment thereof, to bind to or interact with a TLCC target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the TLCC protein to bind to or interact with a TLCC target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+}, diacylglycerol, IP_3, and the like), detecting catalytic/enzymatic activity of the target using an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a TLCC protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the TLCC protein or biologically active portion thereof is determined. Preferred biologically active portions of the TLCC proteins to be used in assays of the present invention include fragments which participate in interactions with non-TLCC molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 2). Binding of the test compound to the TLCC protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the TLCC protein or biologically active portion thereof with a known compound which binds TLCC to form an assay

mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TLCC protein, wherein determining the ability of the test compound to interact with a TLCC protein comprises determining the ability of the test compound to preferentially bind to TLCC or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a TLCC protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TLCC protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a TLCC protein can be accomplished, for example, by determining the ability of the TLCC protein to bind to a TLCC target molecule by one of the methods described above for determining direct binding. Determining the ability of the TLCC protein to bind to a TLCC target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a TLCC protein can be accomplished by determining the ability of the TLCC protein to further modulate the activity of a downstream effector of a TLCC target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a TLCC protein or biologically active portion thereof with a known compound which binds the TLCC protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the TLCC protein, wherein determining the ability of the test compound to interact with the TLCC protein comprises determining the ability of the TLCC protein to preferentially bind to or modulate the activity of a TLCC target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TLCC or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a TLCC protein, or interaction of a TLCC protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ TLCC fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or TLCC protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TLCC binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a TLCC protein or a TLCC target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TLCC protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with TLCC protein or target molecules but which do not interfere with binding of the TLCC protein to its target molecule can be derivatized to the wells of the plate, and unbound target or TLCC protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the TLCC protein

or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TLCC protein or target molecule.

In another embodiment, modulators of TLCC expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TLCC mRNA or protein in the cell is determined. The level of expression of TLCC mRNA or protein in the presence of the candidate compound is compared to the level of expression of TLCC mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TLCC expression based on this comparison. For example, when expression of TLCC mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TLCC mRNA or protein expression. Alternatively, when expression of TLCC mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of TLCC mRNA or protein expression. The level of TLCC mRNA or protein expression in the cells can be determined by methods described herein for detecting TLCC mRNA or protein.

In yet another aspect of the invention, the TLCC proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with TLCC ("TLCC-binding proteins" or "TLCC-bp") and are involved in TLCC activity. Such TLCC-binding proteins are also likely to be involved in the propagation of signals by the TLCC proteins or TLCC targets as, for example, downstream elements of a TLCC-mediated signaling pathway. Alternatively, such TLCC-binding proteins are likely to be TLCC inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a TLCC protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is

fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a TLCC-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the TLCC protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TLCC protein can be confirmed *in vivo, e.g.*, in an animal such as an animal model for a hepatic disorder.

Examples of animal models of hepatic fibrosis include animal models suffering from carbon tetrachloride intoxication, iron and alcohol intoxication, streptococcal cell wall administration, and bile duct ligation, *e.g.*, in rats, as well as mice suffering from schistosomiasis. These animal models are known in the art and are described in, for example, Czaja *et al.* (1989) *J. Cell. Biol.* 108:2477-2482; Manthey *et al.* (1990) *Growth Factors* 4:17-26; Bissell *et al.* (1995) *J. Clin. Invest.* 96:447-455; Tsukamoto *et al.* (1995) *J. Clin. Invest.* 96:620-630; Alcolado *et al.* (1997) *Clin. Sci.* 92:103-112; Cales (1998) *Biomed. and Pharmacother.* 52:259-263. For example, an agent identified as described herein (*e.g.*, a TLCC modulating agent, an antisense TLCC nucleic acid molecule, a TLCC-specific antibody, or a TLCC-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

Animal-based model systems of cardiovascular disease may include, but are not limited to, non-recombinant and engineered transgenic animals.

Non-recombinant animal models for cardiovascular disease may include, for example, genetic models. Such genetic cardiovascular disease models may include, for example, apoB or apoR deficient pigs (Rapacz, *et al.*, 1986, *Science* 234:1573-1577) and Watanabe heritable hyperlipidemic (WHHL) rabbits (Kita *et al.*, 1987, *Proc. Natl. Acad. Sci USA* 84: 5928-5931). Transgenic mouse models in cardiovascular disease and angiogenesis are reviewed in Carmeliet, P. and Collen, D. (2000) J. Pathol. 190:387-405.

Non-recombinant, non-genetic animal models of atherosclerosis may include, for example, pig, rabbit, or rat models in which the animal has been exposed to either chemical wounding through dietary supplementation of LDL, or mechanical wounding through balloon catheter angioplasty. Animal models of cardiovascular disease also include rat myocardial infarction models (described in, for example, Schwarz, ER *et al.* (2000) *J. Am. Coll. Cardiol.* 35:1323-1330) and models of chromic cardiac ischemia in rabbits (described in, for example, Operschall, C *et al.* (2000) *J. Appl. Physiol.* 88:1438-1445).

Models for studying angiogenesis *in vivo* include tumor cell-induced angiogenesis and tumor metastasis (Hoffman, RM (1998-99) *Cancer Metastasis Rev.* 17:271-277; Holash, J *et al.* (1999) *Oncogene* 18:5356-5362; Li, CY *et al.* (2000) *J. Natl Cancer Inst.* 92:143-147), matrix induced angiogenesis (US Patent No. 5,382,514), the disc angiogenesis system (Kowalski, J. *et al.* (1992) *Exp. Mol. Pathol.* 56:1-19), the rodent mesenteric-window angiogenesis assay (Norrby, K (1992) *EXS* 61:282-286), experimental choroidal neovascularization in the rat (Shen, WY *et al.* (1998) *Br. J. Ophthalmol.* 82:1063-1071), and the chick embryo development (Brooks, PC *et al. Methods Mol. Biol.* (1999) 129:257-269) and chick embryo chorioallantoic membrane (CAM) models (McNatt LG *et al.* (1999) *J. Ocul. Pharmacol. Ther.* 15:413-423; Ribatti, D *et al.* (1996) *Int. J. Dev. Biol.* 40:1189-1197), and are reviewed in Ribatti, D and Vacca, A (1999) *Int. J. Biol. Markers* 14:207-213.

Models for studying vascular tone *in vivo* include the rabbit femoral artery model (Luo *et al.* (2000) *J. Clin. Invest.* 106:493-499), eNOS knockout mice (Hannan *et al.* (2000) *J. Surg. Res.* 93:127-132), rat models of cerebral ischemia (Cipolla *et al.* (2000) *Stroke* 31:940-945), the renin-angiotensin mouse system (Cvetkovik *et al.* (2000) *Kidney Int.* 57:863-874), the rat lung transplant model (Suda *et al.* (2000) *J. Thorac.*

Cardiovasc. Surg. 119:297-304), the New Zeland White rabbit model of intracranial hypertension (Richards *et al.* (1999) *Acta Neurochir.* 141:1221-1227), the spontaneously hypertensive (SH) rat neurogenic model of chronic hypertension (Stekiel *et al.* (1999) *Anesthesiology* 91:207-214), the Prague hypertensive rat (PHR) (Vogel *et al.* (1999) *Clin. Sci.* 97:91-98), chronically angiotensin II (Ang II)-infused rats (Pasquie *et al.* (1999) *Hypertension* 33:830-834), Dahl-salt-sensitive rats (Boulanger (1999) *J. Mol. Cell. Cardiol.* 31:39-49), the mouse model of arterial remodeling (Bryant *et al.* (1999) *Circ. Res.* 84:323-328), and the obese Zucker (fa/fa) rat (Golub *et al.* (1998) *Hypertens. Res.* 21:283-288).

Cells that contain and express TLCC gene sequences which encode a TLCC protein, and, further, exhibit cellular phenotypes associated with cardiovascular disease, may be used to identify compounds that exhibit anti-cardiovascular disease activity. Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC#TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as human umbilical vein endothelial cells (HUVECs), human microvascular endothelial cells (HMVEC), and bovine aortic endothelial cells (BAECs); as well as generic mammalian cell lines such as HeLa cells and COS cells, *e.g.*, COS-7 (ATCC# CRL-1651). Further, such cells may include recombinant, transgenic cell lines. For example, the cardiovascular disease animal models of the invention, discussed above, may be used to generate cell lines, containing one or more cell types involved in cardiovascular disease, that can be used as cell culture models for this disorder. While primary cultures derived from the cardiovascular disease transgenic animals of the invention may be utilized, the generation of continuous cell lines is preferred. For examples of techniques which may be used to derive a continuous cell line from the transgenic animals, see Small *et al.*, (1985) *Mol. Cell Biol.* 5:642-648.

Alternatively, cells of a cell type known to be involved in cardiovascular disease may be transfected with sequences capable of increasing or decreasing the amount of TLCC gene expression within the cell. For example, TLCC gene sequences may be introduced into, and overexpressed in, the genome of the cell of interest, or, if endogenous TLCC gene sequences are present, they may be either overexpressed or, alternatively disrupted in order to underexpress or inactivate TLCC gene expression.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the TLCC nucleotide sequences, described herein, can be used to map the location of the TLCC genes on a chromosome. The mapping of the TLCC sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease .

Briefly, TLCC genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TLCC nucleotide sequences. Computer analysis of the TLCC sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the TLCC sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human

chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TLCC nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TLCC sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TLCC gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The TLCC sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the TLCC nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the

sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The TLCC nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from TLCC nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of TLCC Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified

sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the TLCC nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

The TLCC nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such TLCC probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, TLCC primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining TLCC protein and/or nucleic acid expression as well as TLCC activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a

disorder, associated with aberrant or unwanted TLCC expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TLCC protein, nucleic acid expression or activity. For example, mutations in a TLCC gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby phophylactically treat an individual prior to the onset of a disorder characterized by or associated with TLCC protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of TLCC in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of TLCC protein or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting TLCC protein or nucleic acid (*e.g.*, mRNA, or genomic DNA) that encodes TLCC protein such that the presence of TLCC protein or nucleic acid is detected in the biological sample. A preferred agent for detecting TLCC mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TLCC mRNA or genomic DNA. The nucleic acid probe can be, for example, the TLCC nucleic acid set forth in SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to TLCC mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting TLCC protein is an antibody capable of binding to TLCC protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is

directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect TLCC mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of TLCC mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TLCC protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of TLCC genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of TLCC protein include introducing into a subject a labeled anti-TLCC antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting TLCC protein, mRNA, or genomic DNA, such that the presence of TLCC protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of TLCC protein, mRNA or genomic DNA in the control sample with the presence of TLCC protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TLCC in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting TLCC protein or mRNA in a biological sample; means for determining the amount of TLCC in the sample; and means for comparing the amount of TLCC in the sample with a standard. The compound or agent can be packaged in a

suitable container. The kit can further comprise instructions for using the kit to detect TLCC protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder (*e.g.*, hepatic disorder) associated with aberrant or unwanted TLCC expression or activity. As used herein, the term "aberrant" includes a TLCC expression or activity which deviates from the wild type TLCC expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant TLCC expression or activity is intended to include the cases in which a mutation in the TLCC gene causes the TLCC gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional TLCC protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a TLCC substrate, *e.g.*, a non-calcium channel subunit or ligand, or one which interacts with a non-TLCC substrate, *e.g.* a non-calcium channel subunit or ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response, such as cellular proliferation. For example, the term unwanted includes a TLCC expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in TLCC protein activity or nucleic acid expression, such as cardiovascular disease, an endothelial cell disease, a hepatic disorder (*e.g.*, liver fibrosis, hepatitis, liver tumors, cirrhosis of the liver, hemochromatosis, liver parasite induced disorders, alpha-1 antitrypsin deficiency, and autoimmune hepatitis), or a CNS disorder (*e.g.*, a neurodegenerative disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder). Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in TLCC protein activity or nucleic acid expression, such as a hepatic disorder, a CNS disorder, a pain disorder, or a cellular

proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted TLCC expression or activity in which a test sample is obtained from a subject and TLCC protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of TLCC protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted TLCC expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a hepatic disorder, a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted TLCC expression or activity in which a test sample is obtained and TLCC protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of TLCC protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TLCC expression or activity).

The methods of the invention can also be used to detect genetic alterations in a TLCC gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TLCC protein activity or nucleic acid expression, such as a cardiovascular disease, an endothelial cell disorder, a hepatic disorder, a CNS disorder, pain disorder, or a disorder of cellular growth, differentiation, or migration. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TLCC -protein, or the mis-expression of the TLCC gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more

nucleotides from a TLCC gene; 2) an addition of one or more nucleotides to a TLCC gene; 3) a substitution of one or more nucleotides of a TLCC gene, 4) a chromosomal rearrangement of a TLCC gene; 5) an alteration in the level of a messenger RNA transcript of a TLCC gene, 6) aberrant modification of a TLCC gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TLCC gene, 8) a non-wild type level of a TLCC-protein, 9) allelic loss of a TLCC gene, and 10) inappropriate post-translational modification of a TLCC-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a TLCC gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TLCC-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res* .23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a TLCC gene under conditions such that hybridization and amplification of the TLCC-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection

schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a TLCC gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TLCC can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in TLCC can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TLCC gene and detect mutations by comparing the sequence of the sample TLCC with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the

diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the TLCC gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TLCC sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TLCC cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a TLCC sequence, *e.g.*, a wild-type TLCC sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TLCC genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control TLCC nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a TLCC gene.

Furthermore, any cell type or tissue in which TLCC is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a TLCC protein (*e.g.*, the modulation of membrane excitability) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase TLCC gene expression, protein levels, or upregulate TLCC activity, can be monitored in clinical trials of subjects exhibiting decreased TLCC gene expression, protein levels, or downregulated TLCC activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease TLCC gene expression, protein levels, or downregulate TLCC activity, can be monitored in clinical trials of subjects exhibiting increased TLCC

gene expression, protein levels, or upregulated TLCC activity. In such clinical trials, the expression or activity of a TLCC gene, and preferably, other genes that have been implicated in, for example, a TLCC-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including TLCC, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates TLCC activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on TLCC-associated disorders (*e.g.*, disorders characterized by deregulated signaling or membrane excitation), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of TLCC and other genes implicated in the TLCC-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of TLCC or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TLCC protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TLCC protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TLCC protein, mRNA, or genomic DNA in the pre-administration sample with the TLCC protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of TLCC to higher levels than detected,

i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of TLCC to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, TLCC expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted TLCC expression or activity, *e.g.*, a cardiovascular disease, an endothelial cell disorder, a hepatic disorder, a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TLCC molecules of the present invention or TLCC modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

Treatment is defined as the application or administration of a therapeutic agent to a patient, or the application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease, a symptom of disease or a predisposition toward a disease, with the purpose of curing, healing, alleviating, relieving, altering, remedying, ameliorating, improving or affecting the disease, the symptoms of disease or the predisposition toward disease as described herein.

A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted TLCC expression or activity, by administering to the subject a TLCC or an agent which modulates TLCC expression or at least one TLCC activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TLCC expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TLCC aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TLCC aberrancy, for example, a TLCC, TLCC agonist or TLCC antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating TLCC expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a TLCC or agent that modulates one or more of the activities of TLCC protein activity associated with the cell. An agent that modulates TLCC protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a TLCC protein (*e.g.*, a TLCC substrate), a TLCC antibody, a TLCC agonist or antagonist, a peptidomimetic of a TLCC agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more TLCC activities. Examples of such stimulatory agents include active TLCC protein and a nucleic acid molecule encoding TLCC that has been introduced into the cell. In another embodiment, the agent inhibits one or more TLCC activities. Examples of such inhibitory agents include antisense TLCC nucleic acid molecules, anti-TLCC antibodies, and TLCC inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by

culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a TLCC protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) TLCC expression or activity. In another embodiment, the method involves administering a TLCC protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted TLCC expression or activity.

Stimulation of TLCC activity is desirable in situations in which TLCC is abnormally downregulated and/or in which increased TLCC activity is likely to have a beneficial effect. Likewise, inhibition of TLCC activity is desirable in situations in which TLCC is abnormally upregulated and/or in which decreased TLCC activity is likely to have a beneficial effect.

3. Pharmacogenomics

The TLCC molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TLCC activity (*e.g.*, TLCC gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) TLCC-associated disorders (*e.g.*, proliferative disorders) associated with aberrant or unwanted TLCC activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TLCC molecule or TLCC modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TLCC molecule or TLCC modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (*e.g.*, a TLCC protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a TLCC molecule or TLCC modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus

enhance therapeutic or prophylactic efficiency when treating a subject with a TLCC molecule or TLCC modulator, such as a modulator identified by one of the exemplary screening assays described herein.

4. Use of TLCC Molecules as Surrogate Markers

The TLCC molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the TLCC molecules of the invention may be detected, and may be correlated with one or more biological states *in vivo*. For example, the TLCC molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The TLCC molecules of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is

indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker (*e.g.*, a TLCC marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, anti-TLCC antibodies may be employed in an immune-based detection system for a TLCC protein marker, or TLCC-specific radiolabeled probes may be used to detect a TLCC mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The TLCC molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, *e.g.*, McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or protein (*e.g.*, TLCC protein or RNA) for specific tumor

markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in TLCC DNA may correlate TLCC drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

5. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising TLCC sequence information is also provided. As used herein, "TLCC sequence information" refers to any nucleotide and/or amino acid sequence information particular to the TLCC molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said TLCC sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon TLCC sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local

area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the TLCC sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dataprocessor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the TLCC sequence information.

By providing TLCC sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a TLCC- associated disease or disorder or a pre-disposition to a TLCC-associated disease or disorder, wherein the method comprises the steps of determining TLCC sequence information associated with the subject and based on the TLCC sequence information, determining whether the subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC - associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a TLCC-associated disease or disorder or a pre-disposition to a disease associated with a TLCC wherein the method

comprises the steps of determining TLCC sequence information associated with the subject, and based on the TLCC sequence information, determining whether the subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC -associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC -associated disease or disorder associated with TLCC, said method comprising the steps of receiving TLCC sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to TLCC and/or a TLCC-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC information (*e.g.*, sequence information and/or information related thereto), and the acquired information, determining whether the subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC -associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC -associated disease or disorder, said method comprising the steps of receiving information related to TLCC (*e.g.*, sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to TLCC and/or related to a TLCC -associated disease or disorder, and based on one or more of the phenotypic information, the TLCC information, and the acquired information, determining whether the subject has a TLCC -associated disease or disorder or a pre-disposition to a TLCC -associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a TLCC sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be TLCC. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a TLCC-associated disease or disorder, progression of TLCC-associated disease or disorder, and processes, such a cellular transformation associated with the TLCC-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of TLCC expression on the expression of other genes). This provides, for

example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including TLCC) that could serve as a molecular target for diagnosis or therapeutic intervention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and the Sequence Listing, are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN TLCC cDNA

In this example, the identification and characterization of the gene encoding human TLCC (clone Fbh18607) is described.

Isolation of the TLCC cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as TLCC. The entire sequence of the human clone Fbh18607 was determined and found to contain an open reading frame termed human "TLCC."

The nucleotide sequence encoding the human TLCC protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The protein encoded by this nucleic acid comprises about 1130 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone Fbh18607, comprising the coding region of human TLCC, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

Analysis of the Human TLCC Molecule

A BLASTN 2.0 search against the NRN database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC revealed that human TLCC is 97% identical to human STS WI-30695, sequence tagged site (Accession Number G22461) over nucleotides 3874-3605. This search further revealed that human TLCC is homologous to human chromosome 11p15.5 PAC clone pDJ915f1 containing KvLQT1 gene, complete sequence (Accession Number AC003693).

A BLASTN 2.0 search against the dbEST database, using a score of 100 and a word length of 12 (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide sequence of human TLCC revealed that human TLCC is 98% identical to nf99c01.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:928032 (Accession Number AA551759) over nucleotides 3865-3369. This search further revealed that human TLCC is 100% identical to tg78b06.x1Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2114867 (Accession Number AI417040) over nucleotides 3866-3391. This search further revealed that human TLCC is 97% identical to nq58f08.s1 NCI_CGAP_Co9 Homo sapiens cDNA clone IMAGE:1148103 (Accession Number AA633315) over nucleotides 3868-3428. This search further revealed that human TLCC is 98% identical to qp09f02.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1917531 3', mRNA sequence (Accession Number AI344661) over nucleotides 3866-3437). This search further revealed that human TLCC is 97% identical to ah33h08.s1 Soares testis NHT Homo sapiens cDNA clone 1276383 3' (Accession Number AA694490) over nucleotides 3863-3418.

A BLASTN 2.0 search against the PATENT_2/gsnuc database, using a score of 100 and a wordlength of 12, of the nucleotide sequence of human TLCC revealed that human TLCC is 98% identical to human PS112 consensus DNA fragment from gene specific clones (Accession Number V26656) over nucleotides 1509-3900. This search further revealed that human TLCC is 99% identical to full length cDNA sequence of prostate tumor clone J1-17 (Accession Number V61200) over nucleotides 2360-3881. This search further revealed that human TLCC is 99% identical to prostate tumour specific gene clone J1-17 (Accession Number V58585) over nucleotides 2360-3881. This search further revealed that human TLCC is 99% identical to human PS112 5'-EST

DNA fragment (Accession Number V26657) over nucleotides 2614-3900. This search further revealed that human TLCC is 94% identical to 3' cDNA sequence of prostate tumor clone J1-17 (Accession Number V61142) over nucleotides 3204-3755. This search further found that human TLCC is 94% identical to 3' fragment of prostate tumour specific gene J1-17 (Accession Number V58485) over nucleotides 3204-3755. A CLUSTAL W (1.74) alignment of the human TLCC nucleotide sequence with the top hit in this search is provided in Figure 3.

A BLASTN 2.0 search against the PATENT_2/Patent DbPreviewNuc database, using a score of 100 and a wordlength of 12, of the nucleotide sequence of human TLCC revealed that human TLCC is 99% identical to human nucleic acid (Accession Number AC31503 (WO99/46374)) over nucleotides 2339-3886, and 56% identical over nucleotides 3778-3895. This search further revealed that human TLCC is 99% identical to human nucleic acid (Accession Number AC31066 (WO99/46374)) over nucleotides 2621-3170. This search further revealed that human TLCC is 62% identical to 36 secreted proteins (Accession Number AC28066 (WO99/35158)) over nucleotides 2261-3173. This search further revealed that human TLCC is 64% identical to 36 secreted proteins (Accession Number AC28051 (WO99/35158)) over nucleotides 2421-3173.

A BLASTX 2.0 search against the NRP/protot database, using a wordlength of 3, a score of 100, and a BLOSUM62 matrix, of the translated nucleotide sequence of human TLCC revealed that human TLCC is 35% identical to the amino acid sequence of *C. elegans* hypothetical protein CET01H8.1, CEC05C12.3, CEF54D1.5 similar to trp and trp-like proteins [*Homo sapiens*] (Accession Number AB001535) over translated nucleic acid residues 147 to 2018, and 41% identical over translated nucleic acid residues 2205-3470. This search further found that human TLCC is 32% identical to the amino acid sequence of Accession Number Z83117, similarity with Drosophila transient-reporter-potential protein (Swiss Prot accession number P19334); cDNA EST EMBL: D27562 comes from this gene, cDNA EST yk219f12.5 comes from this gene [*Caenorhabditis elegans*] over translated nucleic acid residues 84-1418, 27% identical over translated nucleic acid residues 2190-3368, 30% identical over translated nucleic acid residues 1470-2063, 28% identical over translated nucleic acid residues 3076-3213, 46% identical over translated nucleic acid residues 1613-1651, and 32% identical over translated nucleic acid residues 3705-3839. This search further found that human TLCC

is 33% identical to *Homo sapiens* melastatin I (Accession Number AF071787) over translated nucleic acid residues 2205-3401, 33% identical over translated nucleic acid residues 150-1142, 27% identical over translated nucleic acid residues 1548 to 2405, 48% identical over translated nucleic acid residues 1155-1298, 34% identical over translated nucleic acid residues 3801-3896, 30% identical over translated nucleic acid residues 1261-1380, and 36% identical over translated nucleic acid residues 2451-2516. This search further found that human TLCC is 31% identical to cDNA EST yk308e9.3 comes from this gene; cDNA EST yk308e9.5 comes from this gene; cDNA EST yk318f4.3 comes from this gene; cDNA EST yk318f4.5 comes from this gene; cDNA EST yk398a12.3 comes from this gene, cDNA EST yk398a12.5 comes from this gene (Accession Number Z68333) over translated nucleic acid residues 147-1328, is 23% identical over translated nucleic acid residues 2190-3422, is 31% identical over translated nucleic acid residues 1554-2099, is 34% identical over translated nucleic acid residues 1355-1468, and is 32% identical over translated nucleic acid residues 3225-3338. This search further found that human TLCC is 29% identical to similarity to Worm protein C05C12.3; cDNA EST yk224b10.3 comes from this gene; cDNA EST yk224b10.5 comes from this gene; cDNA EST yk301f12.3 comes from this gene; cDNA EST yk301f12.5 comes from this gene; cDNA EST yk405b7.3 comes from this gene over translated nucleic acid residues 147-2069, is 26% identical over translated nucleic acid residues 2193-2978, and is 34% identical over translated nucleic acid residues 2895-3257. This search further found that human TLCC is 34% identical to *Mus musculus* melastatin (Accession Number AF047714) over translated nucleic acid residues 150-1142, is 48% identical over translated nucleic acid residues 1155-1298, and is 36% identical over translated nucleic acid residues 2427-2516. A CLUSTAL W (1.74) alignment of the translated human TLCC sequence with the top three hits in this search is provided in Figure 4.

A BLASTX 2.0 search against the PATENT_2/gsprot database, using a score of 100, a wordlength of 3 and a BLOSUM62 matrix, of the translated nucleotide sequence of human TLCC revealed that human TLCC is 95% identical to human PS112 protein sequence from gene-specific clones (Accession Number W54425) over translated nucleic acid residues 1509-3524. This search further revealed that human TLCC is 100% identical to amino acid encoded by prostate tumour clone J1-17 (Accession

Number W71868) over translated nucleic acid residues 2580-3524. This search further revealed that human TLCC is 100% identical to prostate tumour specific gene clone J1-17 protein (Accession Number W69384) over translated nucleic acid residues 2580-3524. This search further revealed that human TLCC is 34% identical to prostate-tumour derived antigen #4 (Accession Number Y00931) over translated nucleic acid residues 147-1310, 37% identical over translated nucleic acid residues 2457-3401, 36% identical over translated nucleic acid residues 1554-2018, 46% identical over translated nucleic acid residues 2196-2390, and 38% identical over translated nucleic acid residues 2931-2993. A ClustalW (1.74) alignment of the translated cDNA sequence of human TLCC with the top four hits of this search is provided in Figure 5.

A search was performed against the Memsat database (Figure 6), and correlated with an analysis of the hydrophilicity and surface probability of human TLCC (Figure 2), resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC (SEQ ID NO 2) at about residues 599-619, residues 690-712, residues 784-803, residues 811-831, residues 845-862, and residues 933-957.

A search was also performed against the Prosite database, and resulted in the identification of an N-glycosylation site at residues 143-146, at residues 205-208, and at residues 907-910. The results of the search are set forth in Figure 7.

A search was also performed against the ProDom database (Figure 8) resulting in the identification of a transmembrane calcium channel domain in human TLCC (SEQ ID NO:2) at about residues 783-845. This search further identified significant sequence similarity between the amino acid sequence of human TLCC and human melastatin (Accession Number AAC80000). An alignment (using the GAP program in the GCG software package (Blosum 62 matrix), a gap weight of 12, and a length weight of 4) of the amino acid sequence of human TLCC with human melastatin (Accession Number AAC80000), revealing that human TLCC is 31.739% identical to human melastatin, is set forth in Figure 9.

Tissue Distribution of TLCC mRNA

This example describes the tissue distribution of TLCC mRNA, as was qualitatively determined by Polymerase Chain Reaction (PCR), and quantitatively measured using the Taqman™ procedure.

Using PCR techniques, the human TLCC gene was determined to be predominantly expressed in osteoblasts, with some expression also seen in brain, adipose tissue, breast, colon, all fetal tissues, liver, pituitary, melanocyte, prostate, cervix, muscle, small intestine, megakaryocytes, and aorta, as well as in lymphoma and colon to liver metastases.

Human TLCC expression levels were measured in a variety of tissue and cell samples using the Taqman™ procedure. The Taqman™ procedure is a quantitative, real-time PCR-based approach to detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmplTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA is generated from the samples of interest and serves as the starting materials for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) is included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product.

Using the foregoing Taqman™ procedure, it was determined that TLCC mRNA was expressed at low levels in normal human heart, kidney, lung, and liver. A very marked upregulation was detected in passaged human stellate cells, as well as in human fibrotic livers, although expression was low in quiescent stellate cells. TLCC mRNA was upregulated in human dermal and lung fibroblasts cultured in the presence of TGF-β.

It was determined that the rat orthologue of TLCC was highly increased in all bile duct ligation-induced fibrotic livers tested as compared to control animals. An upregulation was detected in all carbon tetrachloride-induced fibrotic livers as compared to controls. However, there was no significant regulation in the serum-induced fibrotic livers as compared to controls, and no regulation in the cultured rat stellate cells. These data reveal that TLCC is highly regulated in activated stellate cells and in fibrotic livers, being expressed only at low levels in other organs and cell types. These observations suggest that TLCC may play an important role in Ca^{2+}-dependent phenomena (*e.g.*, hepatic cell contractility and proliferation). The functional linkage of TRP channels to inositol triphosphate further suggests that TLCC might be related to key signaling events during stellate cell activation.

EXAMPLE 2: EXPRESSION OF RECOMBINANT TLCC PROTEIN IN BACTERIAL CELLS

In this example, TLCC is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, TLCC is fused to GST and this fusion polypeptide is expressed in *E. coli*, *e.g.*, strain PEB199. Expression of the GST-TLCC fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined...

EXAMPLE 3: EXPRESSION OF RECOMBINANT TLCC PROTEIN IN COS CELLS

To express the TLCC gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TLCC protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the TLCC DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TLCC coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TLCC coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the TLCC gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5α, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the TLCC-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the TLCC polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine

available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the TLCC coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TLCC polypeptide is detected by radiolabelling and immunoprecipitation using a TLCC specific monoclonal antibody.

EXAMPLE 4: EXPRESSION OF TLCC IN BLOOD VESSELS

Reverse Transcriptase PCR (RT-PCR) was performed using the Taqman procedure to detect the presence of RNA transcripts corresponding to human TLCC in mRNA prepared from isolated human vessels or cells cultured from the endothelial vasculature (Figure 10). Bars 1-5 illustrate cultured cells, while bars 6-24 represent isolated human vessels. Significant TLCC expression was detected in vascular smooth muscle cells cultured from human aorta (bars 1 and 2) as well as in endothelial cells cultured from lung microvasculature (bar 3) or umbilical vein. Comparison of bars 4 and 5 indicates that expression of TLCC was downregulated when cultured umbilical vein endothelial cells were treated with human recombinant IL-1β for six hours. Expression of TLCC in several isolated human vessels (bars 7-24) exceeded the expression level of TLCC in human adipose tissue (bar 6) which was included as a control.

EXAMPLE 5: EXPRESSION OF TLCC IN ENDOTHELIAL CELLS DURING LAMINAR SHEAR STRESS

Human umbilical vein endothelial cells (HUVECs) were cultured *in vitro* under standard conditions, described in, for example, U.S. Patent 5,882,925. Experimental cultures were then exposed to laminar shear stress (LSS) conditions.

Cultured HUVEC monolayers were exposed to laminar sheer stress by culturing the cells in a specialized apparatus containing liquid culture medium. Static cultures grown in the same medium served as controls. The *in vitro* LSS treatment at 10 dyns/cm^2 was performed for 24 hours and was designed to simulate the shear stress generated by blood flow in a straight, healthy artery.

The effect of LSS on TLCC expression in endothelial cells was assessed from total RNA prepared from the cells and used to probe clones arrayed on nylon filters. A TLCC clone showed a higher signal when probed with two of the three LSS samples when compared to their static controls, indicating that expression of TLCC is upregulated by laminar shear stress (Figure 11).

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1; and

b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3.

2. An isolated nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2.

3. An isolated nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____.

4. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2.

5. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof;

b) a nucleic acid molecule comprising a fragment of at least 30 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof;

c) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2; and

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2.

6. An isolated nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.

7. An isolated nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.

8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous polypeptide.

9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.

10. The vector of claim 9, which is an expression vector.

11. A host cell transfected with the expression vector of claim 10.

12. A method of producing a polypeptide comprising culturing the host cell of claim 11 in an appropriate culture medium to, thereby, produce the polypeptide.

13. An isolated polypeptide selected from the group consisting of:

a)a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 10 contiguous amino acids of SEQ ID NO:2;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1 or 3 under stringent conditions;

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or 3; and

d) a polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of SEQ ID NO:2.

14. The isolated polypeptide of claim 13 comprising the amino acid sequence of SEQ ID NO:2.

15. The polypeptide of claim 13, further comprising heterologous amino acid sequences.

16. An antibody which selectively binds to a polypeptide of claim 13.

17. A method for detecting the presence of a polypeptide of claim 13 in a sample comprising:

a) contacting the sample with a compound which selectively binds to the polypeptide; and

b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.

18. The method of claim 17, wherein the compound which binds to the polypeptide is an antibody.

19. A kit comprising a compound which selectively binds to a polypeptide of claim 13 and instructions for use.

20. A method for detecting the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in a sample comprising:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to a complement of the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to the complement of the nucleic acid molecule in the sample to thereby detect the presence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.

21. The method of claim 20, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

22. A kit comprising a compound which selectively hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and instructions for use.

23. A method for identifying a compound which binds to a polypeptide of claim 13 comprising:

a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

b) determining whether the polypeptide binds to the test compound.

24. The method of claim 23, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detection of test compound/polypeptide binding;

b) detection of binding using a competition binding assay; and

c) detection of binding using an assay for TLCC activity.

25. A method for modulating the activity of a polypeptide of claim 13 comprising contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

26. A method for identifying a compound which modulates the activity of a polypeptide of claim 13 comprising:

a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

27. The method of claim 26, wherein said activity is modulation of cardiovascular function.

28. The method of claim 26, wherein said activity is modulation of hepatic function.

29. A method for identifying a compound which modulates hepatic function comprising:

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of hepatic function by determining the effect of the test compound on the activity of the polypeptide.

30. A method for identifying a compound which modulates liver fibrosis comprising:

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of liver fibrosis by determining the effect of the test compound on the activity of the polypeptide.

31. A method for treating a subject having a hepatic disorder comprising administering to the subject a TLCC modulator, thereby treating said subject having a hepatic disorder.

32. A method for treating a subject having hepatic disorder comprising administering to the subject a TLCC modulator, wherein the TLCC modulator is the modulator identified by the method of claim 26, thereby treating said subject having a hepatic disorder.

33. The method of claim 31, wherein the TLCC modulator is a small molecule.

34. The method of claim 31, wherein said TLCC modulator is administered in a pharmaceutically acceptable formulation.

35. The method of claim 31, wherein said TLCC modulator is administered using a gene therapy vector.

36. The method of 31, wherein the TLCC modulator is capable of modulating TLCC polypeptide activity.

37. A method for identifying a compound which modulates cardiovascular function comprising:

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of cardiovascular function by determining the effect of the test compound on the activity of the polypeptide.

38. A method for identifying a compound which modulates atherosclerosis comprising:

a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and

b) identifying the compound as a modulator of atherosclerosis by determining the effect of the test compound on the activity of the polypeptide.

39. A method for treating a subject having a cardiovascular disorder comprising administering to the subject a TLCC modulator, thereby treating said subject having a cardiovascular disorder.

40. A method for treating a subject having a cardiovascular disorder comprising administering to the subject a TLCC modulator, wherein the TLCC modulator is the modulator identified by the method of claim 26, thereby treating said subject having a cardiovascular disorder.

41. The method of claim 39, wherein the TLCC modulator is a small molecule.

42. The method of claim 39, wherein said TLCC modulator is administered in a pharmaceutically acceptable formulation.

43. The method of claim 39, wherein said TLCC modulator is administered using a gene therapy vector.

44. The method of 39, wherein the TLCC modulator is capable of modulating TLCC polypeptide activity.

Input file Fbh18607FL.seq; Output File 18607.trans
Sequence length 3900

```
CGGCCCATCTCTCTGGGTCTCTGTCCCTCTCTCTCTGGGTCTCTGTCCCCGTCTCTCTGGGTCTCGGTCCCCGTCTCTC

                                                          M   C   P   Q   F      S
TGGGTCTCTGTCCCCGTCTCTCTGGGTCTCTGTCCCCCTCCCTGTGTGCCCCGCTCCC ATG TGT CCA CAG TTC    15

 L   R   L   S   D   R   T   D   P   A   A   V   Y   S   L   V   T   R   T   W     25
CTC CGG CTC TCT GAC CGA ACG GAT CCA GCT GCA GTT TAT AGT CTG GTC ACA CGC ACA TGG    75

 G   F   R   A   P   N   L   V   V   S   V   L   G   G   S   G   G   P   V   L     45
GGC TTC CGT GCC CCG AAC CTG GTG GTG TCA GTG CTG GGG GGA TCG GGG GGC CCC GTC CTC   135

 Q   T   W   L   Q   D   L   L   R   R   G   L   V   R   A   A   Q   S   T   G     65
CAG ACC TGG CTG CAG GAC CTG CTG CGT CGT GGG CTG GTG CGG GCT GCC CAG AGC ACA GGA   195

 A   W   I   V   T   G   G   L   H   T   G   I   G   R   H   V   G   V   A   V     85
GCC TGG ATT GTC ACT GGG GGT CTG CAC ACG GGC ATC GGC CGG CAT GTT GGT GTG GCT GTA   255

 R   D   H   Q   M   A   S   T   G   G   T   K   V   V   A   M   G   V   A   P    105
CGG GAC CAT CAG ATG GCC AGC ACT GGG GGC ACC AAG GTG GTG GCC ATG GGT GTG GCC CCC   315

 W   G   V   V   R   N   R   D   T   L   I   N   P   K   G   S   F   P   A   R    125
TGG GGT GTG GTC CGG AAT AGA GAC ACC CTC ATC AAC CCC AAG GGC TCG TTC CCT GCG AGG   375

 Y   R   W   R   G   D   P   E   D   G   V   Q   F   P   L   D   Y   N   Y   S    145
TAC CGG TGG CGC GGT GAC CCG GAG GAC GGG GTC CAG TTT CCC CTG GAC TAC AAC TAC TCG   435

 A   F   F   L   V   D   D   G   T   H   G   C   L   G   G   E   N   R   F   R    165
GCC TTC TTC CTG GTG GAC GAC GGC ACA CAC GGC TGC CTG GGG GGC GAG AAC CGC TTC CGC   495

 L   R   L   E   S   Y   I   S   Q   Q   K   T   G   V   G   G   T   G   I   D    185
TTG CGC CTG GAG TCC TAC ATC TCA CAG CAG AAG ACG GGC GTG GGA GGG ACT GGA ATT GAC   555

 I   P   V   L   L   L   L   I   D   G   D   E   K   M   L   T   R   I   E   N    205
ATC CCT GTC CTG CTC CTC CTG ATT GAT GGT GAT GAG AAG ATG TTG ACG CGA ATA GAG AAC   615

 A   T   Q   A   Q   L   P   C   L   L   V   A   G   S   G   G   A   A   D   C    225
GCC ACC CAG GCT CAG CTC CCA TGT CTC CTC GTG GCT GGC TCA GGG GGA GCT GCG GAC TGC   675

 L   A   E   T   L   E   D   T   L   A   P   G   S   G   G   A   R   Q   G   E    245
CTG GCG GAG ACC CTG GAA GAC ACT CTG GCC CCA GGG AGT GGG GGA GCC AGG CAA GGC GAA   735

 A   R   D   R   I   R   R   F   F   P   K   G   D   L   E   V   L   Q   A   Q    265
GCC CGA GAT CGA ATC AGG CGT TTC TTT CCC AAA GGG GAC CTT GAG GTC CTG CAG GCC CAG   795

 V   E   R   I   M   T   R   K   E   L   L   T   V   Y   S   S   E   D   G   S    285
GTG GAG AGG ATT ATG ACC CGG AAG GAG CTC CTG ACA GTC TAT TCT TCT GAG GAT GGG TCT   855

 E   E   F   E   T   I   V   L   K   A   L   V   K   A   C   G   S   S   E   A    305
GAG GAA TTC GAG ACC ATA GTT TTG AAG GCC CTT GTG AAG GCC TGT GGG AGC TCG GAG GCC   915

 S   A   Y   L   D   E   L   R   L   A   V   A   W   N   R   V   D   I   A   Q    325
TCA GCC TAC CTG GAT GAG CTG CGT TTG GCT GTG GCT TGG AAC CGC GTG GAC ATT GCA CAG   975

 S   E   L   F   R   G   D   I   Q   W   R   S   F   H   L   E   A   S   L   M    345
AGT GAA CTC TTT CGG GGG GAC ATC CAA TGG CGG TCC TTC CAT CTC GAA GCT TCC CTC ATG  1035

 D   A   L   L   N   D   R   P   E   F   V   R   L   L   I   S   H   G   L   S    365
GAC GCC CTG CTG AAT GAC CGG CCT GAG TTC GTG CGC TTG CTC ATT TCC CAC GGC CTC AGC  1095
```

FIGURE 1A

```
 L  G  H  F  L  T  P  M  R  L  A  Q  L  Y  S  A  A  P  S  N   385
CTG GGC CAC TTC CTG ACC CCG ATG CGC CTG GCC CAA CTC TAC AGC GCG GCG CCC TCC AAC  1155

 S  L  I  R  N  L  L  D  Q  A  S  H  S  A  G  T  K  A  P  A   405
TCG CTC ATC CGC AAC CTT TTG GAC CAG GCG TCC CAC AGC GCA GGC ACC AAA GCC CCA GCC  1215

 L  K  G  G  A  A  E  L  R  P  P  D  V  G  H  V  L  R  M  L   425
CTA AAA GGG GGA GCT GCG GAG CTC CGG CCC CCT GAC GTG GGG CAT GTG CTG AGG ATG CTG  1275

 L  G  K  M  C  A  P  R  Y  P  S  G  G  A  W  D  P  H  P  G   445
CTG GGG AAG ATG TGC GCG CCG AGG TAC CCC TCC GGG GGC GCC TGG GAC CCT CAC CCA GGC  1335

 Q  G  F  G  E  S  M  Y  L  L  S  D  K  A  T  S  P  L  S  L   465
CAG GGC TTC GGG GAG AGC ATG TAT CTG CTC TCG GAC AAG GCC ACC TCG CCG CTC TCG CTG  1395

 D  A  G  L  G  Q  A  P  W  S  D  L  L  L  W  A  L  L  L  N   485
GAT GCT GGC CTC GGG CAG GCC CCC TGG AGC GAC CTG CTT CTT TGG GCA CTG TTG CTG AAC  1455

 R  A  Q  M  A  M  Y  F  W  E  M  G  S  N  A  V  S  S  A  L   505
AGG GCA CAG ATG GCC ATG TAC TTC TGG GAG ATG GGT TCC AAT GCA GTT TCC TCA GCT CTT  1515

 G  A  C  L  L  L  R  V  M  A  R  L  E  P  D  A  E  E  A  A   525
GGG GCC TGT TTG CTG CTC CGG GTG ATG GCA CGC CTG GAG CCT GAC GCT GAG GAG GCA GCA  1575

 R  R  K  D  L  A  F  K  F  E  G  M  G  V  D  L  F  G  E  C   545
CGG AGG AAA GAC CTG GCG TTC AAG TTT GAG GGG ATG GGC GTT GAC CTC TTT GGC GAG TGC  1635

 Y  R  S  S  E  V  R  A  A  R  L  L  L  R  R  C  P  L  W  G   565
TAT CGC AGC AGT GAG GTG AGG GCT GCC CGC CTC CTC CTC CGT CGC TGC CCG CTC TGG GGG  1695

 D  A  T  C  L  Q  L  A  M  Q  A  D  A  R  A  F  F  A  Q  D   585
GAT GCC ACT TGC CTC CAG CTG GCC ATG CAA GCT GAC GCC CGT GCC TTC TTT GCC CAG GAT  1755

 G  V  Q  S  L  L  T  Q  K  W  W  G  D  M  A  S  T  T  P  I   605
GGG GTA CAG TCT CTG CTG ACA CAG AAG TGG TGG GGA GAT ATG GCC AGC ACT ACA CCC ATC  1815

 W  A  L  V  L  A  F  F  C  P  P  L  I  Y  T  R  L  I  T  F   625
TGG GCC CTG GTT CTC GCC TTC TTT TGC CCT CCA CTC ATC TAC ACC CGC CTC ATC ACC TTC  1875

 R  K  S  E  E  E  P  T  R  E  E  L  E  F  D  M  D  S  V  I   645
AGG AAA TCA GAA GAG GAG CCC ACA CGG GAG GAG CTA GAG TTT GAC ATG GAT AGT GTC ATT  1935

 N  G  E  G  P  V  G  T  A  D  P  A  E  K  T  P  L  G  V  P   665
AAT GGG GAA GGG CCT GTC GGG ACG GCG GAC CCA GCC GAG AAG ACG CCG CTG GGG GTC CCG  1995

 R  Q  S  G  R  P  G  C  C  G  G  R  C  G  G  R  R  C  L  R   685
CGC CAG TCG GGC CGT CCG GGT TGC TGC GGG GGC CGC TGC GGG GGG CGC CGG TGC CTA CGC  2055

 R  W  F  H  F  W  G  A  P  V  T  I  F  M  G  N  V  V  S  Y   705
CGC TGG TTC CAC TTC TGG GGC GCG CCG GTG ACC ATC TTC ATG GGC AAC GTG GTC AGC TAC  2115

 L  L  F  L  L  L  F  S  R  V  L  L  V  D  F  Q  P  A  P  P   725
CTG CTG TTC CTG CTG CTT TTC TCG CGG GTG CTG CTC GTG GAT TTC CAG CCG GCG CCG CCC  2175

 G  S  L  E  L  L  L  Y  F  W  A  F  T  L  L  C  E  E  L  R   745
GGC TCC CTG GAG CTG CTG CTC TAT TTC TGG GCT TTC ACG CTG CTG TGC GAG GAA CTG CGC  2235

 Q  G  L  S  G  G  G  G  S  L  A  S  G  G  P  G  P  G  H  A   765
CAG GGC CTG AGC GGA GGC GGG GGC AGC CTC GCC AGC GGG GGC CCC GGG CCT GGC CAT GCC  2295
```

FIGURE 1B

```
 S  L  S  Q  R  L  R  L  Y  L  A  D  S  W  N  Q  C  D  L  V   785
TCA CTG AGC CAG CGC CTG CGC CTC TAC CTC GCC GAC AGC TGG AAC CAG TGC GAC CTA GTG  2355

 A  L  T  C  F  L  L  G  V  G  C  R  L  T  P  G  L  Y  H  L   805
GCT CTC ACC TGC TTC CTC CTG GGC GTG GGC TGC CGG CTG ACC CCG GGT TTG TAC CAC CTG  2415

 G  R  T  V  L  C  I  D  F  M  V  F  T  V  R  L  L  H  I  F   825
GGC CGC ACT GTC CTC TGC ATC GAC TTC ATG GTT TTC ACG GTG CGG CTG CTT CAC ATC TTC  2475

 T  V  N  K  Q  L  G  P  K  I  V  I  V  S  K  M  M  K  D  V   845
ACG GTC AAC AAA CAG CTG GGG CCC AAG ATC GTC ATC GTG AGC AAG ATG ATG AAG GAC GTG  2535

 F  F  F  L  F  F  L  G  V  W  L  V  A  Y  G  V  A  T  E  G   865
TTC TTC TTC CTC TTC TTC CTC GGC GTG TGG CTG GTA GCC TAT GGC GTG GCC ACG GAG GGG  2595

 L  L  R  P  R  D  S  D  F  P  S  I  L  R  R  V  F  Y  R  P   885
CTC CTG AGG CCA CGG GAC AGT GAC TTC CCA AGT ATC CTG CGC CGC GTC TTC TAC CGT CCC  2655

 Y  L  Q  I  F  G  Q  I  P  Q  E  D  M  D  V  A  L  M  E  H   905
TAC CTG CAG ATC TTC GGG CAG ATT CCC CAG GAG GAC ATG GAC GTG GCC CTC ATG GAG CAC  2715

 S  N  C  S  S  E  P  G  F  W  A  H  P  P  G  A  Q  A  G  T   925
AGC AAC TGC TCG TCG GAG CCC GGC TTC TGG GCA CAC CCT CCT GGG GCC CAG GCG GGC ACC  2775

 C  V  S  Q  Y  A  N  W  L  V  V  L  L  L  V  I  F  L  L  V   945
TGC GTC TCC CAG TAT GCC AAC TGG CTG GTG GTG CTG CTC CTC GTC ATC TTC CTG CTC GTG  2835

 A  N  I  L  L  V  N  L  L  I  A  M  F  S  Y  T  F  G  K  V   965
GCC AAC ATC CTG CTG GTC AAC TTG CTC ATT GCC ATG TTC AGT TAC ACA TTC GGC AAA GTA  2895

 Q  G  N  S  D  L  Y  W  K  A  Q  R  Y  R  L  I  R  E  F  H   985
CAG GGC AAC AGC GAT CTC TAC TGG AAG GCG CAG CGT TAC CGC CTC ATC CGG GAA TTC CAC  2955

 S  R  P  A  L  A  P  P  F  I  V  I  S  H  L  R  L  L  L  R  1005
TCT CGG CCC GCG CTG GCC CCG CCC TTT ATC GTC ATC TCC CAC TTG CGC CTC CTG CTC AGG  3015

 Q  L  C  R  R  P  R  S  P  Q  P  S  S  P  A  L  E  H  F  R  1025
CAA TTG TGC AGG CGA CCC CGG AGC CCC CAG CCG TCC TCC CCG GCC CTC GAG CAT TTC CGG  3075

 V  Y  L  S  K  E  A  E  R  K  L  L  T  W  E  S  V  H  K  E  1045
GTT TAC CTT TCT AAG GAA GCC GAG CGG AAG CTG CTA ACG TGG GAA TCG GTG CAT AAG GAG  3135

 N  F  L  L  A  R  A  R  D  K  R  E  S  D  S  E  R  L  K  R  1065
AAC TTT CTG CTG GCA CGC GCT AGG GAC AAG CGG GAG AGC GAC TCC GAG CGT CTG AAG CGC  3195

 T  S  Q  K  V  D  L  A  L  K  Q  L  G  H  I  R  E  Y  E  Q  1085
ACG TCC CAG AAG GTG GAC TTG GCA CTG AAA CAG CTG GGA CAC ATC CGC GAG TAC GAA CAG  3255

 R  L  K  V  L  E  R  E  V  Q  Q  C  S  R  V  L  G  W  V  A  1105
CGC CTG AAA GTG CTG GAG CGG GAG GTC CAG CAG TGT AGC CGC GTC CTG GGG TGG GTG GCC  3315

 E  A  L  S  R  S  A  L  L  P  P  G  G  P  P  P  P  D  L  P  1125
GAG GCC CTG AGC CGC TCT GCC TTG CTG CCC CCA GGT GGG CCG CCA CCC CCT GAC CTG CCT  3375

 G  S  K  D  *                                                 1130
GGG TCC AAA GAC TGA                                           3390

GCCCTGCTGGCGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGCTCATCTGGGCCTCGGCCCC

CGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCACTGTCAGGACCACCTTTGGGAGTGTCATC
```

FIGURE 1C

CTTACAAACCACAGCATGCCCGGCTCCTCCCAGAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTT

ATCCATCTGGAGGCTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGATTCCTCACACTGGGG

AAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAARRAAAAAAAAAAAAAGGCGG

FIGURE 1D

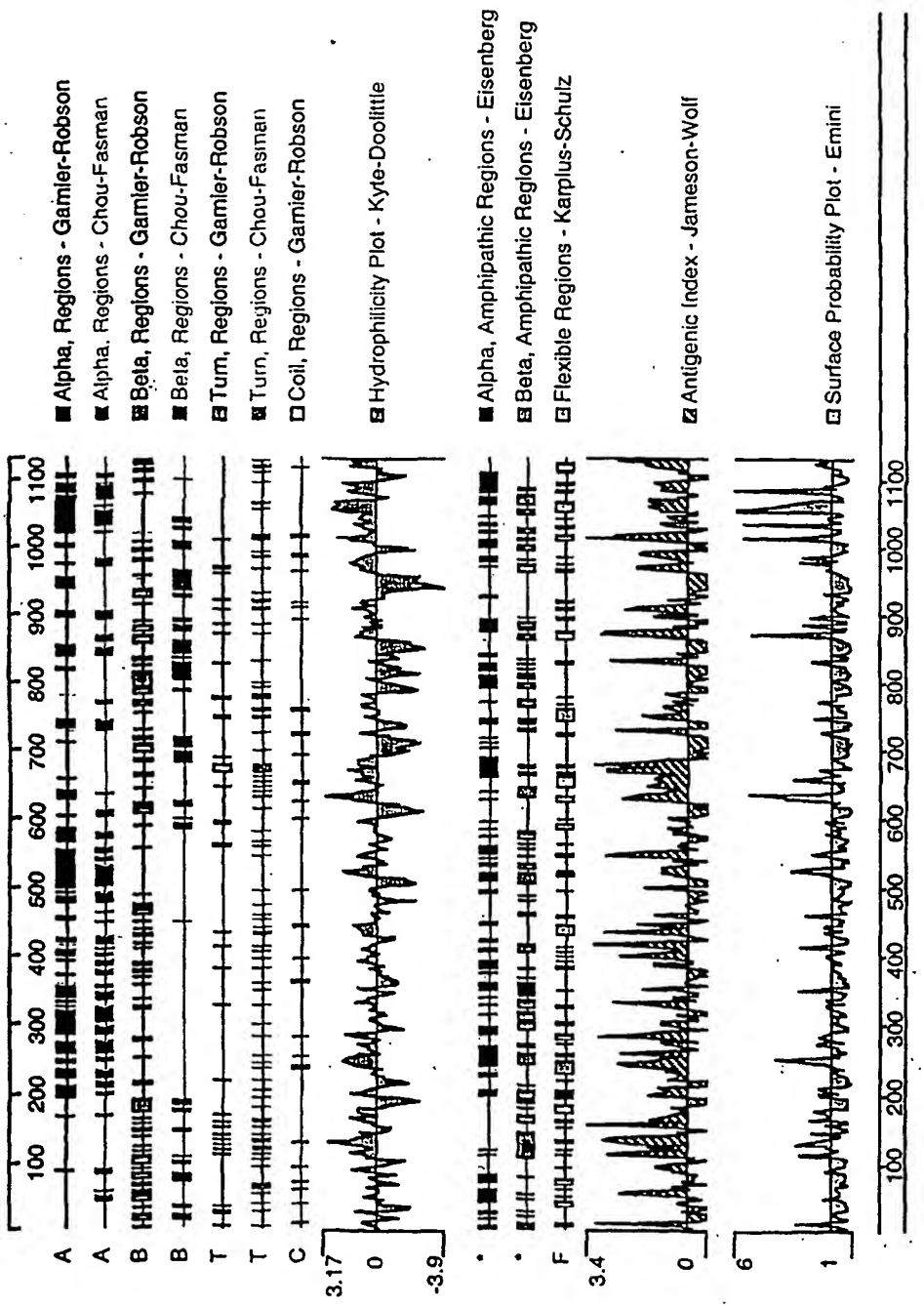

100 200 300 400 500 600 700 800 900 1000 1100
A
A
B
B
T
T
C
Alpha, Regions - Gamier-Robson
Alpha, Regions - Chou-Fasman
Beta, Regions - Gamier-Robson
Beta, Regions - Chou-Fasman
Turn, Regions - Gamier-Robson
Turn, Regions - Chou-Fasman
Coil, Regions - Gamier-Robson
3.17
0
-3.9
Hydrophilicity Plot - Kyte-Doolittle
Alpha, Amphipathic Regions - Eisenberg
Beta, Amphipathic Regions - Eisenberg
Flexible Regions - Karplus-Schulz
3.4
0
Antigenic Index - Jameson-Wolf
6
1
Surface Probability Plot - Emini
100 200 300 400 500 600 700 800 900 1000 1100

FIGURE 2

```
CLUSTAL W (1.74) multiple sequence alignment

Fbh18607FL      CGGCCCATCTCTCTGGGTCTCTGTCCCTCTCTCTCTGGGTCTCTGTCCCCGTCTCTCTGG
V26656          ------------------------------------------------------------

Fbh18607FL      GTCTCGGTCCCCGTCTCTCTGGGTCTCTGTCCCCGTCTCTCTGGGTCTCTGTCCCCCTCC
V26656          ------------------------------------------------------------

Fbh18607FL      CTGTGTGCCCCGCTCCCATGTGTCCACAGTTCCTCCGGCTCTCTGACCGAACGGATCCAG
V26656          ------------------------------------------------------------

Fbh18607FL      CTGCAGTTTATAGTCTGGTCACACGCACATGGGGCTTCCGTGCCCCGAACCTGGTGGTGT
V26656          ------------------------------------------------------------

Fbh18607FL      CAGTGCTGGGGGGATCGGGGGGCCCCGTCCTCCAGACCTGGCTGCAGGACCTGCTGCGTC
V26656          ------------------------------------------------------------

Fbh18607FL      GTGGGCTGGTGCGGGCTGCCCAGAGCACAGGAGCCTGGATTGTCACTGGGGGTCTGCACA
V26656          ------------------------------------------------------------

Fbh18607FL      CGGGCATCGGCCGGCATGTTGGTGTGGCTGTACGGGACCATCAGATGGCCAGCACTGGGG
V26656          ------------------------------------------------------------

Fbh18607FL      GCACCAAGGTGGTGGCCATGGGTGTGGCCCCCTGGGGTGTGGTCCGGAATAGAGACACCC
V26656          ------------------------------------------------------------

Fbh18607FL      TCATCAACCCCAAGGGCTCGTTCCCTGCGAGGTACCGGTGGCGCGGTGACCCGGAGGACG
V26656          ------------------------------------------------------------

Fbh18607FL      GGGTCCAGTTTCCCCTGGACTACAACTACTCGGCCTTCTTCCTGGTGGACGACGGCACAC
V26656          ------------------------------------------------------------

Fbh18607FL      ACGGCTGCCTGGGGGGCGAGAACCGCTTCCGCTTGCGCCTGGAGTCCTACATCTCACAGC
V26656          ------------------------------------------------------------

Fbh18607FL      AGAAGACGGGCGTGGGAGGGACTGGAATTGACATCCCTGTCCTGCTCCTCCTGATTGATG
V26656          ------------------------------------------------------------

Fbh18607FL      GTGATGAGAAGATGTTGACGCGAATAGAGAACGCCACCCAGGCTCAGCTCCCATGTCTCC
V26656          ------------------------------------------------------------
```

FIGURE 3A

```
Fbh18607FL      TCGTGGCTGGCTCAGGGGGAGCTGCGGACTGCCTGGCGGAGACCCTGGAAGACACTCTGG
V26656          ------------------------------------------------------------

Fbh18607FL      CCCCAGGGAGTGGGGGAGCCAGGCAAGGCGAAGCCCGAGATCGAATCAGGCGTTTCTTTC
V26656          ------------------------------------------------------------

Fbh18607FL      CCAAAGGGGACCTTGAGGTCCTGCAGGCCCAGGTGGAGAGGATTATGACCCGGAAGGAGC
V26656          ------------------------------------------------------------

Fbh18607FL      TCCTGACAGTCTATTCTTCTGAGGATGGGTCTGAGGAATTCGAGACCATAGTTTTGAAGG
V26656          ------------------------------------------------------------

Fbh18607FL      CCCTTGTGAAGGCCTGTGGGAGCTCGGAGGCCTCAGCCTACCTGGATGAGCTGCGTTTGG
V26656          ------------------------------------------------------------

Fbh18607FL      CTGTGGCTTGGAACCGCGTGGACATTGCACAGAGTGAACTCTTTCGGGGGGACATCCAAT
V26656          ------------------------------------------------------------

Fbh18607FL      GGCGGTCCTTCCATCTCGAAGCTTCCCTCATGGACGCCCTGCTGAATGACCGGCCTGAGT
V26656          ------------------------------------------------------------

Fbh18607FL      TCGTGCGCTTGCTCATTTCCCACGGCCTCAGCCTGGGCCACTTCCTGACCCCGATGCGCC
V26656          ------------------------------------------------------------

Fbh18607FL      TGGCCCAACTCTACAGCGCGGCGCCCTCCAACTCGCTCATCCGCAACCTTTTGGACCAGG
V26656          ------------------------------------------------------------

Fbh18607FL      CG TCCCACAGCGCAGGCACCAAAGCCCCAGCCCTAAAAGGGGGAGCTGCGGAGCTCCGGC
V26656          ------------------------------------------------------------

Fbh18607FL      CCCCTGACGTGGGGCATGTGCTGAGG ATGCTGCTGGGGAAGATGTGCGCGCCGAGGTACC
V26656          ------------------------------------------------------------

Fbh18607FL      CCTCCGGGGGCGCCTGGGACCCTCACCCAGGCCAGGGCTTCGGGGAGAGC ATGTATCTGC
V26656          ------------------------------------------------------------

Fbh18607FL      TCTCGGACAAGGCCACCTCGCCGCTCTCGCTGGATGCTGGCCTCGGGCAGGCCCCCTGGA
V26656          --------AAGGCCACCTCGCCGCTCTC-CTGGATGCTGGCCTCGG-CAGGCCCCCTGAA
                        ********************* ****************** *********** *

Fbh18607FL      GCGACCTGCTTCTTTGGGCACTGTTGCTGAACAGGGCACAGATGGCCATGTACTTCTGGG
V26656          ----CCTGCTTCTTTGGGCACTGTTGCTGAAACAGGCACAGATGGCCATGTACTTCTGGG
                    ***************************  ***************************
```

FIGURE 3B

```
Fbh18607FL      AGATGGGTTCCAATGCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGG
V26656          AGATGGGTTCCAATGCAGTTTCCTCAGCTCTTGGGGCCTGTTTGCTGCTCCGGGTGATGG
                ************************************************************

Fbh18607FL      CACGCCTGGAGCCTGACGCTGAGGAGGCAGCACGGAGGAAAGACCTGGCGTTCAAGTTTG
V26656          CACGCCTGGAGCCTGACGCTGAGGAGGCAGCACGGAGGAAAGACCTGGCGTTCAAGTTTG
                ************************************************************

Fbh18607FL      AGGGGATGGGCGTTGACCTCTTTGGCGAGTGCTATCGCAGCAGTGAGGTGAGGGCTGCCC
V26656          AGGGGATGGGCGTTGACCTCTTTGGCGAGTGCTATCGCAGCAGTGAGGTGAGGGCTGCCC
                ************************************************************

Fbh18607FL      GCCTCCTCCTCCGTCGCTGCCCGCTCTGGGGGGATGCCACTT-GCCT-CCAGCTGGCCAT
V26656          GCCTCCTCCTCCGTCGYTGCCCGCTCTGGGGGGATGCCACTTTGCCTTCCAGGTGGCCAT
                *************** ************************ **** **** *******

Fbh18607FL      GCAA-GCTGACGCCCGTGCCTTCTTTGCCCAGGATGGGGTACAGTCTCTGCTGACACAGA
V26656          GCAAAGCTGACGSCCSTGMCTTCTTTGCCMAGGATGGGGTACAGTCTCTGCTGACACAGA
                **** ******* ** ** ********** ********************************

Fbh18607FL      AGTGGTGGGGAGATATGGCCAGCACTACACCCATCTGGGCCCTGGTTCTCGCCTTCTTTT
V26656          AGTGGTGGGGAGATATGGCCAGCACTACACCCATCTGGGCCCTGGTTATCGCGTTCTTTT
                ************************************************ **** *******

Fbh18607FL      GCCCTCCACTCATCTACACCCGCCTCATCACCTTCAGGAAATCAGAAGAGGAGCCCACAC
V26656          GCCCTCCACTCATCTACACCCGCCTCATCACCTTCAGGAAATCAGAAGAGGAGCCCACAC
                ************************************************************

Fbh18607FL      GGGAGGAGCTAGAGTTTGACATGGATAGTGTCATTAATGGGGAAGGGCCTGTCGGGACGG
V26656          GGGAGGAGCTAGAGTTTGACATGGATAGTGTCATTAATGGGGAAGGGCCTGTCGGGACGG
                ************************************************************

Fbh18607FL      CGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTCCGGGTTGCT
V26656          CGGACCCAGCCGAGAAGACGCCGCTGGGGGTCCCGCGCCAGTCGGGCCGTCCGGGTTGCT
                ************************************************************

Fbh18607FL      GCGGGGGCCGCTGCGGGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGCGC
V26656          GCGGGGGCCGCTGCGGGGGGCGCCGGTGCCTACGCCGCTGGTTCCACTTCTGGGGCGCGC
                ************************************************************

Fbh18607FL      CGGTGACCATCTTCATGGGCAACGTGGTCAGCTACCTGCTGTTCCTGCTGCTTTTCTCGC
V26656          CGGTGACCATCTTCATGGGCAACGTGGTCAGCTACCTGCTGTTCCTGCTGCTTTTCTCGC
                ************************************************************

Fbh18607FL      GGGTGCTGCTCGTGGATTTCCAGCCGGCGCCGCCCGGCTCCCTGGAGCTGCTGCTCTATT
V26656          GGGTGCTGCTCGTGGATTTCCAGCCGGCGCCGCCCGGCTCCCTGGAGCTGCTGCTCTATT
                ************************************************************

Fbh18607FL      TCTGGGCTTTCACGCTGCTGTGCGAGGAACTGCGCCAGGGCCTGAGCGGAGGCGGGGGCA
V26656          TCTGGGCTTTCACGCTGCTGTGCGAGGAA-TGCGCCAGGGCCTGAGCGGAGGCGGGGGCA
                ********************************* ********************************
```

FIGURE 3C

```
Fbh18607FL      GCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCT
V26656          GCCTCGCCAGCGGGGGCCCCGGGCCTGGCCATGCCTCACTGAGCCAGCGCCTGCGCCTCT
                ************************************************************

Fbh18607FL      ACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTCACCTGCTTCCTCCTGGGCG
V26656          ACCTCGCCGACAGCTGGAACCAGTGCGACCTAGTGGCTCTCACCTGCTTCCTCCTGGGCG
                ************************************************************

Fbh18607FL      TGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCGCACTGTCCTCTGCATCGACT
V26656          TGGGCTGCCGGCTGACCCCGGGTTTGTACCACCTGGGCCGCACTGTCCTCTGCATCGACT
                ************************************************************

Fbh18607FL      TCATGGTTTTCACGGTGCGGCTGCTTCACATCTTCACGGTCAACAAACAGCTGGGGCCCA
V26656          TCATGGTTTTCACGGTGCGGCTGCTTCACATCTTCACGGTCAACAAACAGCTGGGGCCCA
                ************************************************************

Fbh18607FL      AGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTCTTCCTCGGCG
V26656          AGATCGTCATCGTGAGCAAGATGATGAAGGACGTGTTCTTCTTCCTCTTCTTCCTCGGCG
                ************************************************************

Fbh18607FL      TGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGCTCCTGAGGCCACGGGACAGTGACT
V26656          TGTGGCTGGTAGCCTATGGCGTGGCCACGGAGGGGCTCCTGAGGCCACGGGACAGTGACT
                ************************************************************

Fbh18607FL      TCCCAAGTATCCTGCGCCGCGTCTTCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTC
V26656          TCCCAAGTATCCTGCGCCGCGTCTTCTACCGTCCCTACCTGCAGATCTTCGGGCAGATTC
                ************************************************************

Fbh18607FL      CCCAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCT
V26656          CCCAGGAGGACATGGACGTGGCCCTCATGGAGCACAGCAACTGCTCGTCGGAGCCCGGCT
                ************************************************************

Fbh18607FL      TCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGC
V26656          TCTGGGCACACCCTCCTGGGGCCCAGGCGGGCACCTGCGTCTCCCAGTATGCCAACTGGC
                ************************************************************

Fbh18607FL      TGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGTGGCCAACATCCTGCTGGTCAACTTGC
V26656          TGGTGGTGCTGCTCCTCGTCATCTTCCTGCTCGTGGCCAACATCCTGCTGGTCAACTTGC
                ************************************************************

Fbh18607FL      TCATTGCCATGTTCAGTTACACATTCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGA
V26656          TCATTGCCATGTTCAGTTACACATTCGGCAAAGTACAGGGCAACAGCGATCTCTACTGGA
                ************************************************************

Fbh18607FL      AGGCGCAGCGTTACCGCCTCATCCGGGAATTCCACTCTCGGCCCGCGCTGGCCCCGCCCT
V26656          AGGCGCAG-GTTACCGCCTCATCCGGGAATTCCACTCTCGGCCCGCGCTGGCCCCGCCCT
                ******** ***************************************************

Fbh18607FL      TTATCGTCATCTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGACCCCGGAGCC
V26656          TTATCGTCATCTCCCACTTGCGCCTCCTGCTCAGGCAATTGTGCAGGCGACCCCGGAGCC
                ************************************************************
```

FIGURE 3D

```
Fbh18607FL      CCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGC
V26656          CCCAGCCGTCCTCCCCGGCCCTCGAGCATTTCCGGGTTTACCTTTCTAAGGAAGCCGAGC
                ************************************************************

Fbh18607FL      GGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTGCTGGCACGCGCTAGGG
V26656          GGAAGCTGCTAACGTGGGAATCGGTGCATAAGGAGAACTTTCTGCTGGCACGCGCTAGGG
                ************************************************************

Fbh18607FL      ACAAGCGGGAGAGCGACTCCGAGCGTCTGAAGCGCACGTCCCAGAAGGTGGACTTGGCAC
V26656          ACAAGCGGGAGAGCGACTCCGAGCGTCTGAAGCGCACGTCCCAGAAGGTGGACTTGGCAC
                ************************************************************

Fbh18607FL      TGAAACAGCTGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGG
V26656          TGAAACAGCTGGGACACATCCGCGAGTACGAACAGCGCCTGAAAGTGCTGGAGCGGGAGG
                ************************************************************

Fbh18607FL      TCCAGCAGTGTAGCCGCGTCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGC
V26656          TC CAGCAGTGTAGCCGCGTCCTGGGGTGGGTGGCCGAGGCCCTGAGCCGCTCTGCCTTGC
                ************************************************************

Fbh18607FL      TGCCCCCAGGTGGGCCGCCACCCCCTGACCTGCCTGGGTCCAAAGACTGAGCCCTGCTGG
V26656          TGCCCCCAGGTGGGCCGCCACCCCCT GACCTGCCTGGGTCCAAAGACTGAGCCCTGCTGG
                ************************************************************

Fbh18607FL      CGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGCTCATCTGGGC
V26656          CGGACTTCAAGGAGAAGCCCCCACAGGGGATTTTGCTCCTAGAGTAAGGC TCATCTGGGC
                ************************************************************

Fbh18607FL      CTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCAC
V26656          CTCGGCCCCCGCACCTGGTGGCCTTGTCCTTGAGGTGAGCCCCATGTCCATCTGGGCCAC
                ************************************************************

Fbh18607FL      TGTCAGGACCACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGGCTCCTCCCA
V26656          TGTCAGGACCACCTTTGGGAGTGTCATCCTTACAAACCACAGCATGCCCGGCTCCTCCCA
                ************************************************************

Fbh18607FL      GAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTATCCATCTGGAGG
V26656          GAACCAGTCCCAGCCTGGGAGGATCAAGGCCTGGATCCCGGGCCGTTATCCATCTGGAGG
                ************************************************************

Fbh18607FL      CTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGATTCCTCACA
V26656          CTGCAGGGTCCTTGGGGTAACAGGGACCACAGACCCCTCACCACTCACAGATTCCTCACA
                ************************************************************

Fbh18607FL      CTGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG
V26656          CTGGGGAAATAAAGCCATTTCAGAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAG-GG
                ********************************************  **********  **

Fbh18607FL      CGG-------
V26656          CGGCCGCGGT
                ***
```

FIGURE 3E

CLUSTAL W (1.74) multiple sequence alignment

```
3928756         ----------------------------------------------------------M
18607           ----------------------------------------------------------
3243075         ----------------------------------------------------------
3878614         MNLCYRRHRYASSPEVWCTMESDELGVTRYLQSKGGDQVPPTSTTTGGAGGDGNAVPTTS

3928756         EPSALRKAGSEQEEGFEGLPRRVTDLGMVSNLRRSNSSLFKSWRLQCPFGNNDKQESLSS
18607           ----------------------------------------------------------
3243075         ----------------------------------------------------------
3878614         QAQAQTFNSGRQTTGMSSGDRLNEDVSATAN--SAQLVLPTPLFNQMRFTESNMSLNRHN

3928756         WIPENIKKKECVYFVESSKLSDAGKVVCQCGYTHEQH--LEEAT----------------
18607           ----------------------------------------------------------
3243075         ----------------------------------------------------------
3878614         WVRETFTRRECSRFIASSR--DLHK--CGCGRTRDAHRNIPELTSEFLRQKRSVAALEQQ

3928756         -------------KPHTFQGT--QWDPKKHVQEMPTDAFGDIVFTG-LSQKVKKYVRVSQ
18607           ----------------------------------------------MCPQFLRLSD
3243075         ----------------------------------------------MYIRVSY
3878614         RSISNVNDDINTQNMYTKRGANEKWSLRKHTVSLATNAFGQVEFQGGPHPYKAQYVRVNF
                                                                ::*:.

3928756         DTPSSVIYHLMTQHWGLDVPNLLISVTGGAKNFNMKPRLKSIFRRGLVKVAQTTGAWIIT
18607           RTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRRGLVRAAQSTGAWIVT
3243075         DTKPDSLLHLMVKDWQLELPKLLISVHGGLQNFEMQPKLKQVFGKGLIKAAMTTGAWIFT
3878614         DTEPAYIMSLFEHVWQISPPRLIITVHGGTSNFDLQPKLARVFRKGLLKAASTTGAWIIT
                 * .  :  *. : * :  *.*:::* **  .  ::. *  :: :**::.* :*****.*

3928756         GGSHTGVMKQVGEAVRDFSLSSSYKEGELITIGVATWGTVHRREGLIHPTGSFPAEYILD
18607           GGLHTGIGRHVGVAVRDHQMASTGG-TKVVAMGVAPWGVVRNRDTLINPKGSFPARYRWR
3243075         GGVSTGVISHVGDALKDHSSKSRGR---VCAIGIAPWGIVENKEDLVG--KDVTRVYQTM
3878614         SGCDTGVVKHVAAALEGAQSAQRNK---IVCIGIAPWGLLKKREDFIG--QDKTVPYYPS
                .*  **:  :*. *:.. .  .      :  :*:*.** :..:: ::    . .  *

3928756         EDGQ-GNLTCLDSNHSHFILVDDGTHGQYGVEIPLRTRLEKFISEQTKERGGVAIKIPIV
18607           GDPEDGVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQQKTGVGGTGIDIPVL
3243075         SNPL-SKLSVLNNSHTHFILADNGTLGKYGAEVKLRRLLEKHISLQKINTR-LGQGVPLV
3878614         SSKG--RFTGLNNRHSYFLLVDNGTVGRYGAEVILRKRLEMYISQKQKIFG-GTRSVPVV
                 .        *:  :: *:*.*:** *  * *  :*  ** .** :          :*::

3928756         CVVLEGGPGTLHTIDNATTNG--TPCVVVEGSGRVADVIAQVANLPVSDITISLIQQKLS
18607           LLLIDGDEKMLTRIENATQAQ--LPCLLVAGSGGAADCLAETLEDTLAPGSGGARQGE-A
3243075         GLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGIINESLREQL
3878614         CVVLEGGSCTIRSVLDYVTNVPRVPVVVCDGSGRAADLLAFAHQNVTEDGLLPDDIRRQV
                 ::::*.   :  : :         * ::  *** .:* :: . :              .

3928756         VFFQEMFETFTESRIVEWTKKIQDIVRRRQLLTVFREGKDGQQDVDVAILQALLKASRSQ
18607           R--DRIRRFFPKGDLEVLQAQVERIMTRKELLTVYSS-EDGSEEFETIVLKALVKACGSS
3243075         LVTIQKTFNYNKAQSHQLFAIIMECMKKKELVTVFRMGSEGQQDIEMAILTALLKGTNVS
3878614         LLLVETTFGCSEAAAHRLLHELTVCAQHKNLLTIFRLGEQGEHDVDHAILTALLKGQNLS
                    .      :.        :     :::*:*::   .:*..:.:  :* **:*.   .
```

FIGURE 4A

```
3928756         DHFGHENWDHQLKLAVAWNRVDIARSEIFMDEWQWKPS----------
18607           E---ASAYLDELRLAVAWNRVDIAQSELFRGDIQWRSF----------
3243075         AP-------DQLSLALAWNRVDIARSQIFVFGPHWTPLGSLAPPTDSKATEKEKKPPMAT
3878614         AA-------DQLALALAWNRVDIARSDVFAMGHEWPQA----------
                      .:* **:********:*::*     .*

3928756         ----------------------------------DLHPTMTAALISNKPEFVKLFLEN
18607           ----------------------------------HLEASLMDALLNDRPEFVRLLISH
3243075         TKGGRGKGKGKKKGKVKEEVEEETDPRKIELLNWVNALEQAMLDALVLDRVDFVKLLIEN
3878614         ----------------------------------ALHNAMMEALIHDRVDFVRLLLEQ
                                                   *. ::  **: :: :**:*::.:

3928756         GVQLKEFVTWDTLLYLYEN-LDPSCLFHSKL-QKVLVEDPERPACAPAAP----------
18607           GLSLGHFLTPMRLAQLYSA-APSNSLIRNLLDQASHSAGTKAPALKGGAA----------
3243075         GVNMQHFLTIPRLEELYNTRLGPPNTLHLLVRDVKKSNLPPDYHISLIDIGLVLEYLMGG
3878614         GINMQKFLTISRLDELYNTDKGPPNTLFYIVRDVVRVRQGYRFKLPDIGL--VIEKLMGN
                *:.: ..*:*   *  **.    .   :   : :

3928756         RLQMHHVAQVLR------------------------------------------------
18607           ELRPPDVGHVLR------------------------------------------------
3243075         AYRCNYTRKNFR------------------------------------------------
3878614         SYQCSYTTSEFRDKYKQRMKRVKHAQKKAMGVFSSRPSRTGSGIASRQSTEGMGGVGGGS
                  :   .   :*

3928756         ---ELLGDFTQ----PLYP----------RPRHNDRLRL----------LLPVPHVKLNV
18607           ---MLLGKMCA----PRYP----------SGGAWD----------PHPGQGFG-
3243075         ---TLYNNLFG----PKRP----------KALKLLG-ME----------DDEPPAKGKKK
3878614         SVAGVFGNSFGNQDPPLDPHVNRSALSGSRALSNHILWRSAFRGNFPANPMRPPNLGDSR
                    : ..       *  *                                  *

3928756         ----QGVSLRSLYKRS--SGHVTFTMDPIRDLLIWAIVQNRRELAGIIWAQSQDCIAAAL
18607           ----ESMYLLSDKATSPLSLDAGLGQAPWSDLLLWALLLNRAQMAMYFWEMGSNAVSSAL
3243075         ----KKKKKEEEIDIDVDDPAVSRFQYPFHELMVWAVLMKRQKMAVFLWQRGEESMAKAL
3878614         DCGSEFDEELSLTSASDGSQTEPDFRYPYSELMIWAVLTKRQDMAMCMWQHGEEAMAKAL
                    :     .   .  .          *  :*::**:: :* .:*  :*  ..:.:: **

3928756         ACSKILKELSKEEED----TDSSEEMLALAEEYEHRAIGVFTECYRKDEERAQKLLTRVS
18607           GACLLLRVMARLEPD----AEEAARRKDLAFKFEGMGVDLFGECYRSSEVRAARLLLRRC
3243075         VACKLYKAMAHESSESDLVDDISQDLDNNSKDFGQLALELLDQSYKHDEQIAMKLLTYEL
3878614         VACRLYKSLATEAAEDYLEVEICEELKKYAEEFRILSLELLDHCYHVDDAQTLQLLTYEL
                  .. : : ::    :     : .      : .:   .: ::..*: .:  : :**

3928756         EAWGKTTCLQLALEAKDMKFVSHGGIQAFLTKVWWG--QLSVDNGLWRVTLCMLAFPLLL
18607           PLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWG--DMASTTPIWALVLAFFCPPLIY
3243075         KNWSNSTCLKLAVAAKHRDFIAHTCSQMLLTDMWMGRLRMRKNPGLKVIMGILLPPTILF
3878614         SNWSNETCLALAVIVNNKHFLAHPCCQILLADLWHGGLRMRTHSNIKVVLGLICPPFIQM
                  *.. *** **: ..   *.::   * :*:. * *   :     :  :   :    :

3928756         TGLISFR-----------------------------------------EKR----------
18607           TRLITFRKS---------------------------------------EEEPTR-------
3243075         LEFRTYDDFSYQ-----------------------------------TSKENEDGK-------
3878614         LEFKTREELLNQPQTAAEHQNDMNYSSSSSSSSSSSSSSSSSDSSSFEDDDDENNAHNHD
                  : :                                          . . *.

3928756         ------------------------------LQ---D-----------------VG
```

FIGURE 4B

```
18607          ------------------------------EELEFDMDS----------VINGEGPVG
3243075        ------------------------------EKEEENTDA----------N----ADAG
3878614        QKRTRKTSQGSAQSLNITSLFHSRRRKAKKNEKCDRETDASACEAGNRQIQNGGLTAEYG
                                                :   *                  *

3928756        T---------------PAAR-----------------------------------AR----
18607          TA-------------DPAEKT-------------------PLGVPRQS------GRPGCC
3243075        SR-------------KGDEEN----------------------EHKK------QRSIPI
3878614        TFGESNGVSPPPPYMRANSRSRYNNRSDMSKTSSVIFGSDPNLSKLQKSNITSTDRPNPM
               :                                                        *

3928756        ----------------AFFTAPVVVFHLNILSYFAFLCLFAYVLMVDFQPVP-SWCECAI
18607          GGRCGG--RRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGSLELLL
3243075        G-----------TKICEFYNAPIVKFWFYTISYLGYLLLFNYVILVRMDGWP-SLQEWIV
3878614        EQFQGTRKIKMRRRFYEFYSAPISTFWSWTISFILFITFFTYTLLVKTPPRP-TVIEYIL
                             *: **:  *    :*:: :: :*  .::*     *    *  :

3928756        YLWLFSLVCEEMRQLFYD-----------PDECGLMKKAALYFSDFWNKLDVGAILLFVA
18607          YFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTCFLL
3243075        ISYIVSLALEKIREILMS------------EPGKLSQKIKVWLQEYWNITDLVAISTFMI
3878614        IAYVAAFGLEQVRKIIMS------------DAKPFYEKIRTYVCSFWNCVTILAIIFYIV
                 :  ::  *::*: : .                : ::   :. . **   : *:  ::

3928756        GLTCRLIPATLY-PGRVILSLDFILFCLRLMHIFTISKTLGPKIIIVKRMMKDVFFFLFL
18607          GVGCRLTPGLYH-LGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFF
3243075        GAILRLQNQPYMGYGRVIYCVDIIFWYIRVLDIFGVNKYLGPYVMMIGKMMIDMLYFVVI
3878614        GFFMRCFGSVAY--GRVILACDSVLWTMKLLDYMSVHPKLGPYVTMAGKMIQNMSYIIVM
               *   *          **.: . * :.: :::: : :   *** : :  :*: :: ::::.:

3928756        LAVWVVSFGVAKQAILIHNERRVDWLFRGAVYHSYLTIFGQIPG-YIDGVNFNPEHCSPN
18607          LGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVALMEHSNCS--
3243075        MLVVLMSFGVARQAILHPEEKPSWKLARNIFYMPYWMIYGEVFADQIDLYAMEINPPCG-
3878614        LVVTLLSFGLARQSITYPDETWHWILVRNIFLKPYFMLYGEVYADEIDTCGDEAWDQHLE
               : * ::::*:* :.:    :     : *  .  .*  ::*::    :*    :

3928756        GTDP-YKPKCPESDATQQRPAFPEWLTVLLLCLYLLFTNILLLNLLIAMFNYTFQQVQEH
18607          -SEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKVQGN
3243075        --EN--LYDEEGKRLPPCIPG--AWLTPALMACYLLVANILLVNLLIAVFNNTFFEVKSI
3878614        NGGP--VILGNGTTGLSCVPG--YWIPPLLMTFFLLIANILLMSMLIAIFNHIFDATDEM
                                .    *:   *:  :**.:****:.:***:*.  *   ..

3928756        TDQIWKFQRHDLIEEYHGRPAAPPPFILLSHLQLFIKRVVLK----TPAKRHKQLKN---
18607          SDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRP--RSPQPSSPALEHF--
3243075        SNQVWKFQRYQLIMTFHDRPVLPPPMIILSHIYIIIMRLSGRC--RKKREGDQEERDRGL
3878614        SQQIWLFQRYKQVMEYESTPFLPPPLTPLYHGVLILQFVRTRLSCSKSQERNPILLLKIA
               ::  *  **:  :  :.. *  .**:  : *  :::  :  :    .

3928756        --KLEKNEEAALLSWEIYLKENYLQNRQFQQKQRPEQKIEDISNKVDAM-----------
18607          RVYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVD-------------
3243075        KLFLSDEELKRLHEFEEQCVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERE
3878614        ELFLDNDQIEKLHDFEEDCMEDLARQKLNEKNTSNEQRILRADIRTDQILNR--------
                  *.:    *  :*    ::    :  ::.   .:::  . :.:

3928756        -----VDLLDLDPLKR--SGSMEQRLASLEE---QVAQTARALHWIVRTLRASG-FSSEA
18607          --------LALKQLGH--IREYEQRLKVLER---EVQQCSRVLGWVAEALSRS------A
3243075        TFMKTSLQTVDLRLAQ--LEELSNRMVNALE---NLAGIDRSDLIQARSRASSE-CEATY
```

FIGURE 4C

```
3878614          ----LIDLQAKESMGRDVINDVESRLASVEKAQNEILECVRALLNQNNAPTAIGRCFSPS
                             : :    . ..*:    .   ::    *      .:

3928756          DVPT--------------------LASQKAAEEPDAE----PGGRKKTEEPGDSYHVNAR
18607            LLP-------------------------------------PGGPPPPDLPG-SKD----
3243075          LLRQSSINSADGYSLYRYHFNGEELLFEDTSLSTSPC----TGVRKKTCSFRIKEEKDVK
3878614          PDP---------------------LVETANGTPGPLLLKLPGTDPILEEKD--HDSGEN
                                                        .*

3928756          HLLYPNCP---VTRFPVPNEKVPWETEFLIYDPP----------FYTAERKDAAAMDPMG
18607            ------------------------------------------------------------
3243075          THLVPECQNSLHLSLGTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDSKK
3878614          SNSLPPGRIRRNRTATICGGYVSEERNMMLLSPK---------PSDVSGIPQQRLMSVTS

3928756          DTLEPLSTIQYNVVDGLRDRRSFHGPYTVQAG----LPLNPMGRTGLRGRGSLSCFGPNH
18607            ------------------------------------------------------------
3243075          EETISPSLNKTDVIHGQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACK
3878614          MDPLPLPLAKLSTMS-IRRRHEEYTSITDSIA-----IRHPERRIRNNRSNSSEHDESAV

3928756          TLYPMVTRWRRNEDGAICRKSIKKMLEVLVVKLPLSEHWALPGGSREPGEMLPRKLKRIL
18607            ------------------------------------------------------------
3243075          TMKSRSFVYSRGRKLVGGVNQDVEYSSITDQQLTTEWQCQVQKITRSHSTDIPYIVSEAA
3878614          DSEGGGNVTSSPRKRS---TRDLRMTPSSQVEESTSRDQIFEIDHPEHEEDEAQADCELT

3928756          RQEHWPSFENLLKCGMEVYKGYMDDPRN--TDNAWIETVAVSVHFQDQNDVELNRLNSNL
18607            ------------------------------------------------------------
3243075          VQAEQKEQFADMQDEHHVAEAIPRIPRLSLTITDRNGMENLLSVKPDQTLGFPSLRSKSL
3878614          DVITEEEDEEEDDEEDDSHERHHIHPRR--KSSRQNRQPS-HTLETDLSEGEEVDPLDVL

3928756          HACDSGASIRWQVVDRR--------IPLYANHKTLLQKAAAEFGAHY---
18607            ------------------------------------------------
3243075          HGHPRNVKSIQGKLDRSGHASSVSSLVIVSGMTAEEKKVKKEKASTETEC
3878614          K--MKELPIIHQILNEEEQAGAPHSTPVIASPSSSRADLTSQKCSDV---
```

FIGURE 4D

CLUSTAL W (1.74) multiple sequence alignment

```
W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           ------------------------------------------------------------
Y00931          MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCK

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           ------------------------------------------MCPQFLRLSDRTDPAAV
Y00931          CGYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEIL

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           YSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRRGLVRAAQSTGAWIVTGGLHTGI
Y00931          YELLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSR-LIYIAQSKGAWILTGGTHYGL

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           GRHVGVAVRDHQMASTGGTKVVAMGVAPWGVVRNRDTLI---NPKGSFPARYRWRGDPED
Y00931          TKYIGEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRD

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           GVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQQKTGVGGTGIDIPVLLLLID
Y00931          P-LYILDNNHTHLLLVDNGCHGHPTVEAKLRNQLEKHISERTIQDSNYGGKIPIVCFAQG

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           GDEKMLTRIENATQAQLPCLLVAGSGGAADCLAETLEDTLAPGSGGARQGEARDRIRRF-
Y00931          GGKETLKAINTSIKNKIPCVVVEGSGRIADVIASLVEVEDAPTSSAVKEKLVRFLPRTVS

W54425          ------------------------------------------------------------
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           -FPKGDLEVLQAQVERIMTRKELLTVYSSEDGSEEFET-IVLKALVKACGSSEAS--AYL
Y00931          RLSEEETESWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWN
```

FIGURE 5A

```
W54425  ------------------------------------------------------------
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   DELRLAVAWNRVDIAQSELFRGDIQWRSFHLEASLMDALLNDRPEFVRLLISHGLSLGHF
Y00931  GQLKLLLEWNQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKF

W54425  ------------------------------------------------------------
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   LTPMRLAQLYSAAPSNSLIRNLLDQASHSAGTKAPALKGGAAELRPPDVGHVLRMLLGKM
Y00931  LTHDVLTELFSNHFSTLVYRN-LQIAKNSYN------------------DALLTFVWKL

W54425  ------------------------------KATSPLSWMLASAGPLNLLLWALLLKQAQM
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   CAPRYPSGGAWDPHPGQGFGESMYLLSDKATSPLSLDAGLGQAPWSDLLLWALLLNRAQM
Y00931  VAN-----------FRRGFRKEDRNGRDEMDIELHDVSPITRHPLQALFIWAILQNKKEL

W54425  AMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGMGVDLFGECYRSS
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   AMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGMGVDLFGECYRSS
Y00931  SKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELANEYETRAVELFTECYSSD

W54425  EVRAARLLLRRCPLWGDA---TLPSRWPCKADFFADGVQSLLTQKWWGDMASTTPIWALV
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   EVRAARLLLRRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMASTTPIWALV
Y00931  EDLAEQLLVYSCEAWGGSNCLELAVEATDQHFTAQ-GVQNFLSKQWYGEISRDTKNWKII

W54425  IAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPLGVPRQSG
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   LAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPLGVPRQSG
Y00931  LCLFIIPLVGCGFVSFRK--------K-------------------PVDKH---------

W54425  RPGCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGSLE
W69384  ------------------------------------------------------------
W71868  ------------------------------------------------------------
18607   RPGCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLFSRVLLVDFQPAPPGSLE
Y00931  -----------KKLLWYYVAFFTSPFVVFSWNVVFYIAFLLLFAYVLLMDFHSVPHPP-E
```

FIGURE 5B

```
W54425          LLLYFWAFTLLCEMR-QGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTC
W69384          ------------------------------------------------------------
W71868          ------------------------------------------------------------
18607           LLLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWNQCDLVALTC
Y00931          LVLYSLVFVLFCDEVRQWYVNGVN--------------------YFTDLWNVMDTLGLFY


W54425          FLLGVGCRLTPG---LYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVF
W69384          ----------------------------MVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVF
W71868          ----------------------------MVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVF
18607           FLLGVGCRLTPG---LYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVF
Y00931          FIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRMLIDVF
                                        ::**:**:*****.::*****::::.:*: ***

W54425          FFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVALMEHS
W69384          FFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVALMEHS
W71868          FFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVALMEHS
18607           FFLFFLGVWLVAYGVATEGLLRPRDSDFPSIIRRVFYRPYLQIFGQIPQEDMDVALMEHS
Y00931          FFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVP-SDVDGTTYDFA
                ****::.**:**:*** :*:** .:. :  *:* *:*.*** :***:* .*:* :  :.:

W54425          NCSSEPGFWAHPPGAQAGT-CVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKV
W69384          NCSSEPGFWAHPPGAQAGT-CVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKV
W71868          NCSSEPGFWAHPPGAQAGT-CVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKV
18607           NCSSEPGFWAHPPGAQAGT-CVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYTFGKV
Y00931          HCTFT-GNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGTV
                :*:   *  ::*  .: .   :.:::.:*:.: *: *::* :********:***.**.*.*

W54425          QGNSDLYWKAQ-YRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFR
W69384          QGNSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFR
W71868          QGNSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFR
18607           QGNSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPALEHFR
Y00931          QENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKK-CFKCCCKEKN---MESS-
                * *.*  ** * * *::*: **  :. ****:::: ::::: * :  . : .   :*

W54425          VYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLALKQL-GHIREYE
W69384          VYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLALKQL-GHIREYE
W71868          VYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLALKQL-GHIREYE
18607           VYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLALKQL-GHIREYE
Y00931          VCCFKNEDNETLAWEGVMKENYLVK-INTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIA
                *   *: :.: *:**.* ***:*:   . * :. **.:::  :::*  *::* * ::*

W54425          QRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
W69384          QRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
W71868          QRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
18607           QRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD
Y00931          NKIK----------------------------------------
                :::*
```

FIGURE 5C

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
32	49	ins-->out	1.7
599	619	out-->ins	3.7
690	712	ins-->out	4.1
784	803	out-->ins	2.0
845	862	ins-->out	4.4
933	957	out-->ins	7.4

>18607
MCPQFLRLSDRTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQDLLRRGLVRA
AQSTGAWIVTGGLHTGIGRHVGVAVRDHQMASTGGTKVVAMGVAPWGVVRNRDTLINPKG
SFPARYRWRGDPEDGVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQQKTGVG
GTGIDIPVLLLLIDGDEKMLTRIENATQAQLPCLLVAGSGGAADCLAETLEDTLAPGSGG
ARQGEARDRIRRFFPKGDLEVLQAQVERIMTRKELLTVYSSEDGSEEFETIVLKALVKAC
GSSEASAYLDELRLAVAWNRVDIAQSELFRGDIQWRSFHLEASLMDALLNDRPEPVRLLI
SHGLSLGHFLTPHRLAQLYSAAPSNSLIRNLLDQASHSAGTKAPALKGGAAELRPPDVGH
VLRMLLGKMCAPRYPSGGAWDPHPGQGFGESMYLLSDKATSPLSLDAGLGQAPWSDLLLW
ALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAEEAARRKDLAFKFEGMGVD
LFGECYRSSEVRAARLLLRRCPLWGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDMA
STTPIWALVLAFFCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKT
PLGVPRQSGRPGCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLPLLLFSRVLLVDF
QPAPPGSLELLLYFWAFTLLCEELRQGLSGGGGSLASGGPGPGHASLSQRLRLYLADSWN
QCDLVALTCFLLGVGCRLTPGLYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGPKIVIVSK
MMKDVFFFLFFLGVWLVAYGVATEGLLRPRDSDFPSILRRVFYRFYLQIFGQIPQEDMDV
ALMEHSNCSSEPGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSY
TFGKVQGNSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRPRSPQPSSPA
LEHFRVYLSKEAERKLLTWESVHKENFLLARARDKRESDSERLKRTSQKVDLALKQLGHI
RETEQRLKVLEREVQQCSRVLGWVAEALSRSALLPPGGPPPPDLPGSKD

FIGURE 6

Prosite Pattern Matches for 18607

Prosite version: Release 12.2 of February 1995

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 143 NYSA 146

Query: 205 NATQ 208

Query: 907 NCSS 910

FIGURE 7

View Prodom 17189

>17189 p99.2 (4) O75560(1) Q17652(1) Q93971(1) // PROTEIN CHROMOSOME TRANSMEMBRANE MELASTATIN C05C12.3 T01H8.5 I F54D1.5 IV
Length = 311

Score = 200 (75.5 bits), Expect = 1.7e-13, P = 1.7e-13
Identities = 68/298 (22%), Positives = 124/298 (41%)

```
Query:   819 VRLLHIFTVNKQLGPKIVIVSKMMKDVXXXXXXXGVWLVAYGVATEGLLRPRDSDFPSIL 878
             ++L    +V+ +LGP I + +KM+ ++         V L+A+GVA + +  P   D+  IL
Sbjct:     1 MKLFDYLSVHPKLGPYINMAAKMIWNMCYICVLLLVTLMAFGVARQAITYPDVEDWHWIL 60

Query:   879 -RRVFYRPYLQIFGQIPQEDMDVALME----HSNCSSEPGFWAHPPGAQAGTCVSQYANW 933
              R +FY+PY  ++G++  +++D    E    H   ++ P +  +  G      C+  Y  W
Sbjct:    61 VRNIFYKPYFMLYGEVYADEIDTCGDEIWPPHGE-NNGPIYMEN--GTTGPPCIPGY--W 115

Query:   934 XXXXXXXXXXXXXXXXXXXXXXXXMFSYTFGKVQGNSDLYWKAQRYRLIREFHSRPALAPP 993
                                     +F+  F +    S   W  QRY+ I E+H  P L PP
Sbjct:   116 IPPLLMTCFLLVANILLMNLLIAVFNNIFDETIEMSKQIWLFQRYQQIMEYHDTPFLPPP 175

Query:   994 FIVISHXXXXXXXXXXXXXXXXXXXXAL----EHFRVYLSKEAERKLLTWESVHKENFLL 1049
             F ++ H                          +++LS +  ++L  +E    E+
Sbjct:   176 FTILYHVYWIIQYLKGRLSCSKKQERKQFRRERSLKLFLSDDEMKRLHDFEEDCVEDMTR 235

Query:  1050 ARARDKRESDSERLKRTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVAEA 1107
              +  +K  S+ ER+ RTS++ D     +L  + + E  +K    E+Q     L  +  A
Sbjct:   236 EKEBEKLSSNDERILRTSERTDNICNRLHDLNQKEFMMK---DEIQDVETRLAHIENA 290
```

View Prodom 13079

>13079 p99.2 (5) // PROTEIN MELASTATIN CHROMOSOME TRANSMEMBRANE C05C12.3 T01H8.5 I F54D1.5 IV
Length = 357

Score = 442 (160.7 bits), Expect = 5.2e-41, P = 5.2e-41
Identities = 106/338 (31%), Positives = 181/338 (53%)

```
Query:     5 FLRLSDRTDPAAVYSLVTRTWGFRAPNDXXXXXXXXXXXXXXQTWLQDLLRRGLVRAAQST 64
             ++R+S  T+PA +  L+ + W    P                Q  L  + R+GL++AAQ+T
Sbjct:     2 YVRISYDTEPADIMHLMEKVWQLEPPRLIITVHGGLSNFDLQPKLARVFRKGLLKAAQTT 61

Query:    65 GAWIVTGGLHTGIGRHVGVAVRDHQMASTGGTKVVAMGVAPWGVVRNRDTLINPKGSPPA 124
             GAWI+T GL TG+ +HV  A+ DH  +++    ++VA+G+APWG+++ ++  I    +
Sbjct:    62 GAWIITSGLDTGVVKHVASALHDHGNSASMRNRIVAIGIAPWGMIKRKEDFIGQDET--V 119

Query:   125 RYRWRGDPEDGVQFPLDYNYSAFFLVDDGTHGCLGGENRFRLRLESYISQQKTGVGGTGI 184
              Y+       +      L+  +S F L D+GT G  G E   R RLE +I+QQ+      G
Sbjct:   120 YYQTHSFSVNSRLSVLNDRHSYFPLLADNGTVGRYGAEIHLRKRLEKHIAQQQCNTR-KGR 178

Query:   185 DIPVLLLLIDGDEKMLTRIEN-ATQA-QLPCLLVAGSGGAADCLAETLEDTLAPG--SGG 240
              +PV+ ++++G    +  + +  T   ++P ++  GSG AAD L+    +   G  S
Sbjct:   179 SVPVVCVVLEGGASTINMVHDYVTNVPRIPVIVCDGSGRAADILSFAHQYVNEDGILSDD 238

Query:   241 ARQGEARDRIRRFF--PKGDLEVLQAQVERIMTRKELLTVYS-SEDGSEEFETIVLKALV 297
              R+ +  + I++ F    + D    L  ++    M RK+LLT++    E+G E+ + ++L AL+
Sbjct:   239 IRE-QLLNLIKKTFGYSEADAHQLFRKITECMKRKDLLTIFRLGEEGEEDVDHVILTALL 297

Query:   298 KACGSSEASAYLDELRLAVAWNRVDIAQSELFRGDIQW 335
             K    S      D+L LA+AWNRVDIA+S++F    +W
Sbjct:   298 KGQNLSPP----DQLALALAWNRVDIARSZIFANGHEW 331
```

FIGURE 8A

View Prodom 16694 [Blast] [Showing region] [Scale]

```
>16694 p99.2 (4) O75560(1) Q17652(1) Q93971(1)  // PROTEIN CHROMOSOME
       TRANSMEMBRANE MELASTATIN C05C12.3 T01H8.5 I F54D1.5 IV
       Length = 204

 Score = 175 (66.7 bits), Expect = 2.0e-12, P = 2.0e-12
 Identities = 49/174 (28%), Positives = 76/174 (43%)

Query:   471 QAPWSDXXXXXXXXXNRAQMAMYFWEMGSNAVSSALGACLLLRVMARLEPDAE--EAARRK 528
             Q P+ D          R +MA   W+ G  A++  L AC L + MA  E   E  E    +
Sbjct:    22 QYPFHDLMIWAVLTKRQKMAECMWQHGEEAMAKCLVACRLYKAMAH-EASEEYLEVDISQ 80

Query:   529 DL---AFKFEGMGVDLFGECYRSSEVXXXXXXXXXCPLWGDATCLQLAMQADARAFFAQD 585
             DL    + +F   + ++L  +CY+               P WG+ TCL LA+ A+ R F A
Sbjct:    81 DLDBNSKEFRELALELLDQCYKHDHDQTLRLLTYELPNWGNNTCLSLAVLANHRDFLAHP 140

Query:   586 GVQSLLTQKWWGDMAST-TPIWALVLAFFCPPLIYTRLITFRKSEE---EPTRE 635
               Q LL   W G +     P   ++    CPP I    + F+  ++   +P++E
Sbjct:   141 CCQMLLADLWHGGLRMRKNPNIKVITGLICPPTIL--FLEFKTKDDFSYQPSKE 192
```

View Prodom 12379 [Blast] [Showing region] [Scale]

```
>12379 p99.2 (5)  // PROTEIN MELASTATIN CHROMOSOME TRANSMEMBRANE T01H8.5 I
       C05C12.3 F54D1.5 IV
       Length = 121

 Score = 96 (38.9 bits), Expect = 0.00054, P = 0.00054
 Identities = 31/91 (34%), Positives = 47/91 (51%)

Query:   340 LEASLMDALLNDRPEFVRLLISHGLSLGHFLTPHRLAQLYSA---APSNSLIRNLLDQAS 396
             L  ++MDAL  DR +FV LL+ +G+++  FLT  RL  LY+      P N+L R  +
Sbjct:    15 LHNAMMDALYWDRVDFVHLLLENGVNMQKFLTINRLEHLYNTDDKGPPNTL-RTWVRDVD 73

Query:   397 HSAGTKAPALKGGAAELRPPDVGHVLRMLLG 427
              S     P          ++ PD+G V+  L+G
Sbjct:    74 KS--NVDPHY-----HIKLPDIGLVVEKLMG 97
```

View Prodom 2328 [Blast] [Showing region] [Scale]

```
>2328 p99.2 (23)  // PROTEIN CHANNEL CALCIUM RECEPTOR IONIC TRANSMEMBRANE ION
       TRANSPORT ENTRY TRANSIENT
       Length = 272

 Score = 71 (30.1 bits), Expect = 43., P = 1.0
 Identities = 18/64 (28%), Positives = 33/64 (51%)

Query:   783 DLVALTCFLLGVGCRLTPGLYHLGRTVLCIDFMVFTVRLLHIFTVNKQLGP-KIVIVSKM 841
             +L  L C    + C+       +   +  I  +   +RL++IFT NK LGP +I + ++M
Sbjct:    52 NLFLLICIPFRLACKHEPEPSLIAEALFAIANVFSYLRLIYIFTANKHLGPLQISLGTRM 111

Query:   842 MKDV 845
             + D+
Sbjct:   112 IVDI 115
```

FIGURE 8B

```
 GAP of: 18607orf.pep  check: 9844  from: 1  to: 1130

18607orf

 to: hmelastatin.pep  check: 9477  from: 1  to: 1533

hMELASTATIN 3243075 in GenPept

 Symbol comparison table:
/ddm_local/gcg/gcg_9.1/gcgcore/data/rundata/blosum62.cmp
 CompCheck: 6430

        Gap Weight:     12      Average Match:  2.912
     Length Weight:      4   Average Mismatch: -2.003

           Quality:    684             Length:   1576
             Ratio:  0.605               Gaps:     28
Percent Similarity: 43.698   Percent Identity: 31.739

       Match display thresholds for the alignment(s):
                  | = IDENTITY
                  : =   2
                  . =   1

18607orf.pep x hmelastatin.pep

       .         .         .         .         .
   1 MCPQFLRLSDRTDPAAVYSLVTRTWGFRAPNLVVSVLGGSGGPVLQTWLQ 50
       ::|.|  | | ..  |. : |    | |.:|| ||     :|  |.
   1 ...MYIRVSYDTKPDSLLHLMVKDWQLELPKLLISVHGGLQNFEMQPKLK 47
       .         .         .         .         .
  51 DLLRRGLVRAAQSTGAWIVTGGLHTGIGRHVGVAVRDHQMASTGGTKVVA 100
     .  :||::|| .||||| |||. ||:  ||| |.:||   | |  :| |
  48 QVFGKGLIKAAMTTGAWIFTGGVSTGVISHVGDALKDHSSKSRG..RVCA 95
       .         .         .         .         .
 101 MGVAPWGVVRNRDTLINPKGSFPARYRWRGDPEDGVQFPLDYNYSAFFLV 150
     -|:||||:| |:: |:         |.   .|   .   |. .:. | |
  96 IGIAPWGIVENKEDLVGK..DVTRVYQTMSNPLSKLSV.LNNSHTHFILA 142
       .         .         .         .         .
 151 DDGTHGCLGGENRFRLRLESYISQQKTGVGGTGIDIPVLLLLIDGDEKML 200
     |.|| |  | | : |  || :|| ||       |  :|.. |.::|   ..
 143 DNGTLGKYGAEVKLRRLLEKHISLQKINT.RLGQGVPLVGLVVEGGPNVV 191
       .         .         .         .         .
 201 TRIENATQAQ..LPCLLVAGSGGAADCLAETLEDTLAPG.SGGARQGEAR 247
     . :    | :  :| .:  ||| |.| |.   .    |    . . :
 192 SIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGIINESLREQLL 241
       .         .         .         .         .
 248 DRIRRFF..PKGDLEVLQAQVERIMTRKELLTVY.SSEDGSEEFETIVLK 294
       |.: |   | .   | | :   | :|||.||:    :| :: |  :|
 242 VTIQKTFNYNKAQSHQLFAIIMECMKKKELVTVFRMGSEGQQDIEMAILT 291
       .         .         .         .         .
 295 ALVKACGSSEASAYLDELRLAVAWNRVDIAQSELFRGDIQWRSF..... 338
     ||.|    |    . |:| ||.||||||||.|::|     |
 292 ALLKGTNVSAP.....DQLSLALAWNRVDIARSQIFVFGPHWTPLGSLAPP 337
```

FIGURE 9A

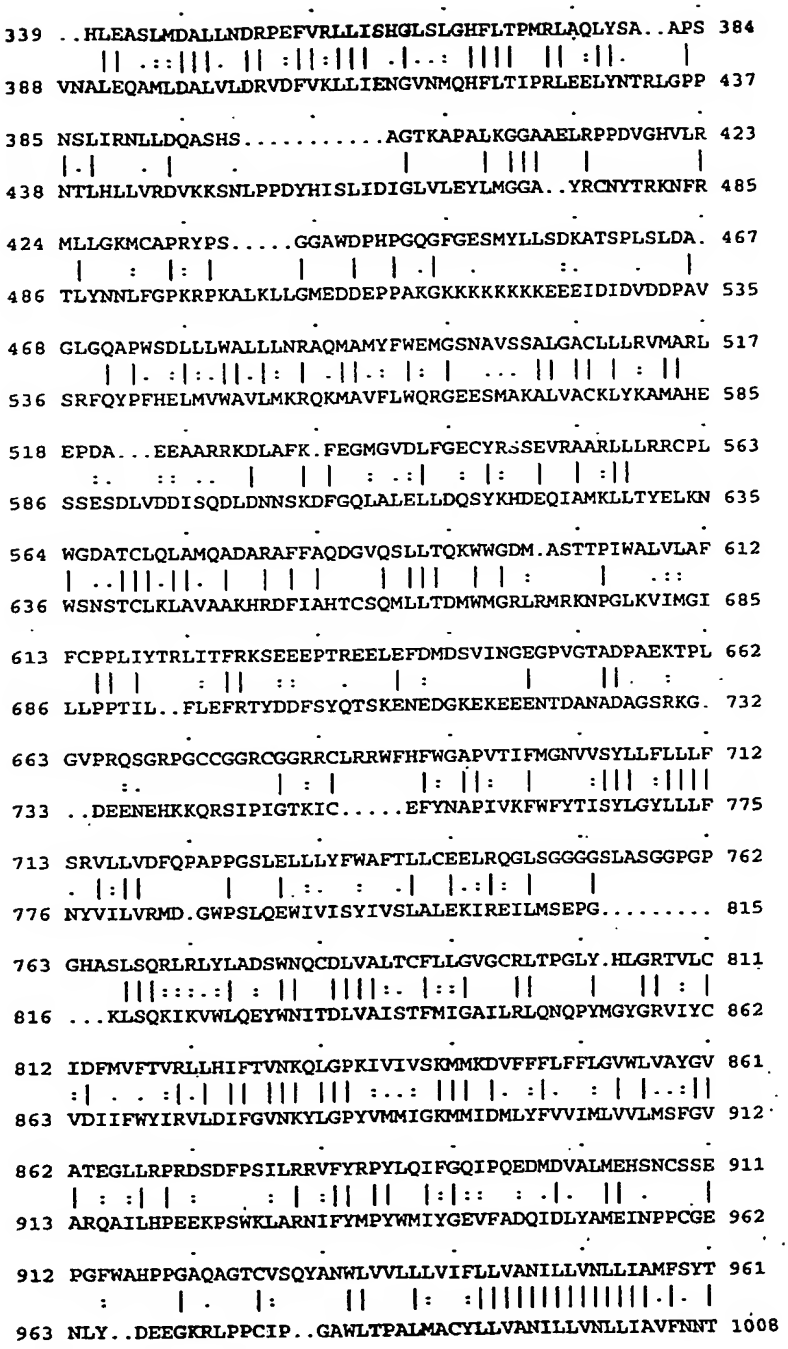

339 ..HLEASLMDALLNDRPEFVRLLISHGLSLGHFLTPMRLAQLYSA..APS 384
388 VNALEQAMLDALVLDRVDFVKLLIENGVNMQHFLTIPRLEELYNTRLGPP 437
385 NSLIRNLLDQASHS.....AGTKAPALKGGAAELRPPDVGHVLR 423
438 NTLHLLVRDVKKSNLPPDYHISLIDIGLVLEYLMGGA..YRCNYTRKNFR 485
424 MLLGKMCAPRYPS.....GGAWDPHPGQGFGESMYLLSDKATSPLSLDA. 467
486 TLYNNLFGPKRPKALKLLGMEDDEPPAKGKKKKKKKKEEEIDIDVDDPAV 535
468 GLGQAPWSDLLLWALLLNRAQMAMYFWEMGSNAVSSALGACLLLRVMARL 517
536 SRFQYPFHELMVWAVLMKRQKMAVFLWQRGEESMAKALVACKLYKAMAHE 585
518 EPDA...EEAARRKDLAFK.FEGMGVDLFGECYRSSEVRAARLLLRRCPL 563
586 SSESDLVDDISQDLDNNSKDFGQLALELLDQSYKHDEQIAMKLLTYELKN 635
564 WGDATCLQLAMQADARAFFAQDGVQSLLTQKWWGDM.ASTTPIWALVLAF 612
636 WSNSTCLKLAVAAKHRDFIAHTCSQMLLTDMWMGRLRMRKNPGLKVIMGI 685
613 FCPPLIYTRLITFRKSEEEPTREELEFDMDSVINGEGPVGTADPAEKTPL 662
686 LLPPTIL..FLEFRTYDDFSYQTSKENEDGKEKEEENTDANADAGSRKG. 732
663 GVPRQSGRPGCCGGRCGGRRCLRRWFHFWGAPVTIFMGNVVSYLLFLLLF 712
733 ..DEENEHKKQRSIPIGTKIC.....EFYNAPIVKFWFYTISYLGYLLLF 775
713 SRVLLVDFQPAPPGSLELLLYFWAFTLLCEELRQGLSGGGGSLASGGPGP 762
776 NYVILVRMD.GWPSLQEWIVISYIVSLALEKIREILMSEPG..... 815
763 GHASLSQRLRLYLADSWNQCDLVALTCFLLGVGCRLTPGLY.HLGRTVLC 811
816 ...KLSQKIKVWLQEYWNITDLVAISTFMIGAILRLQNQPYMGYGRVIYC 862
812 IDFMVFTVRLLHIFTVNKQLGPKIVIVSKMMKDVFFFLFFLGVWLVAYGV 861
863 VDIIFWYIRVLDIFGVNKYLGPYVMMIGKMMIDMLYFVVIMLVVLMSFGV 912
862 ATEGLLRPRDSDFPSILRRVFYRPYLQIFGQIPQEDMDVALMEHSNCSSE 911
913 ARQAILHPEEKPSWKLARNIFYMPYWMIYGEVFADQIDLYAMEINPPCGE 962
912 PGFWAHPPGAQAGTCVSQYANWLVVLLLVIFLLVANILLVNLLIAMFSYT 961
963 NLY..DEEGKRLPPCIP..GAWLTPALMACYLLVANILLVNLLIAVFNNT 1008

FIGURE 9B

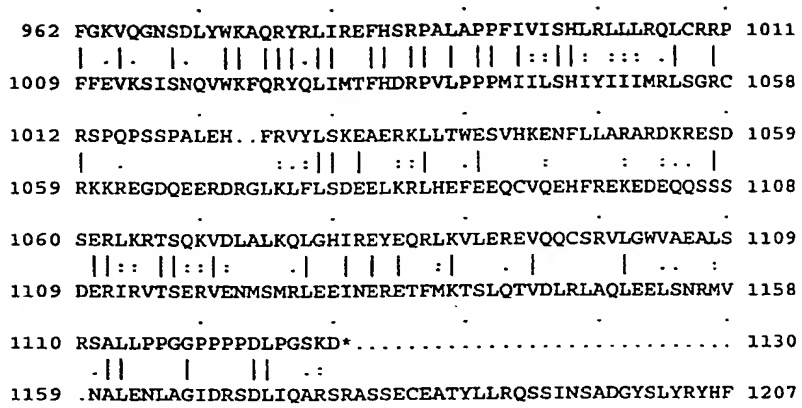
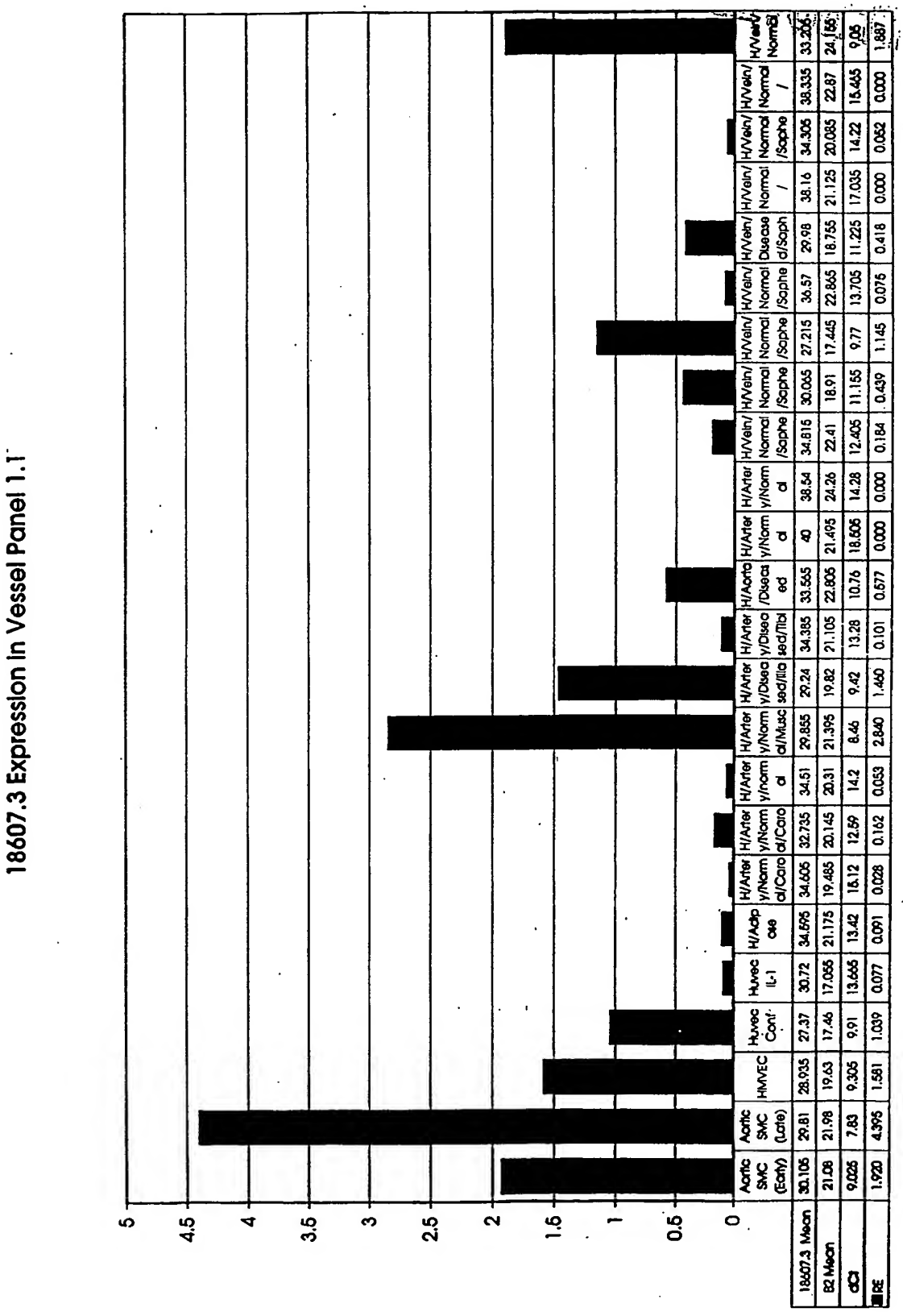

962 FGKVQGNSDLYWKAQRYRLIREFHSRPALAPPFIVISHLRLLLRQLCRRP 1011
1009 FFEVKSISNQVWKFQRYQLIMTFHDRPVLPPPMIILSHIYIIIMRLSGRC 1058
1012 RSPQPSSPALEH..FRVYLSKEAERKLLTWESVHKENFLLARARDKRESD 1059
1059 RKKREGDQEERDRGLKLFLSDEELKRLHEFEEQCVQEHFREKEDEQQSSS 1108
1060 SERLKRTSQKVDLALKQLGHIREYEQRLKVLEREVQQCSRVLGWVAEALS 1109
1109 DERIRVTSERVENMSMRLEEINERETFMKTSLQTVDLRLAQLEELSNRMV 1158
1110 RSALLPPGGPPPPDLPGSKD*..... 1130
1159 .NALENLAGIDRSDLIQARSRASSECEATYLLRQSSINSADGYSLYRYHF 1207

FIGURE 9C



	Aortic SMC (Early)	Aortic SMC (Late)	HMVEC	Huvec Cont	Huvec IL-1	H/Adip ose	H/Aorta y/Normal/Carotid	H/Aorta y/Normal/Carotid	H/Aorta y/normal	H/Aorta y/Normal/Musc ul	H/Aorta y/Disease/Fibroblast	H/Aorta /Diseased	H/Aorta y/Normal	H/Aorta y/Normal	H/Vein/Normal/Saphe	H/Vein/Normal/Saphe	H/Vein/Normal/Saphe	H/Vein/Disease/Saphe	H/Vein/Normal/Saphe	H/Vein/Normal/Saphe	H/Vein/Normal
18607.3 Mean	30.105	29.81	28.935	27.37	30.72	34.595	34.606	32.735	34.51	29.855	29.24	34.385	33.535	40	38.54	34.815	30.065	27.215	36.57	29.98	38.335
B2 Mean	21.08	21.98	19.63	17.46	17.055	21.175	19.485	20.145	20.31	21.395	19.82	21.105	22.805	21.495	24.26	22.41	18.91	17.445	22.865	18.755	22.87
dCt	9.025	7.83	9.305	9.91	13.665	13.42	15.12	12.59	14.2	8.46	9.42	13.28	10.76	18.606	14.28	12.405	11.155	9.77	13.705	11.225	15.465
RE	1.920	4.395	1.581	1.039	0.077	0.091	0.028	0.162	0.053	2.840	1.460	0.101	0.677	0.000	0.000	0.184	0.439	1.145	0.075	0.418	0.000	0.052	0.000	1.887

Figure 10

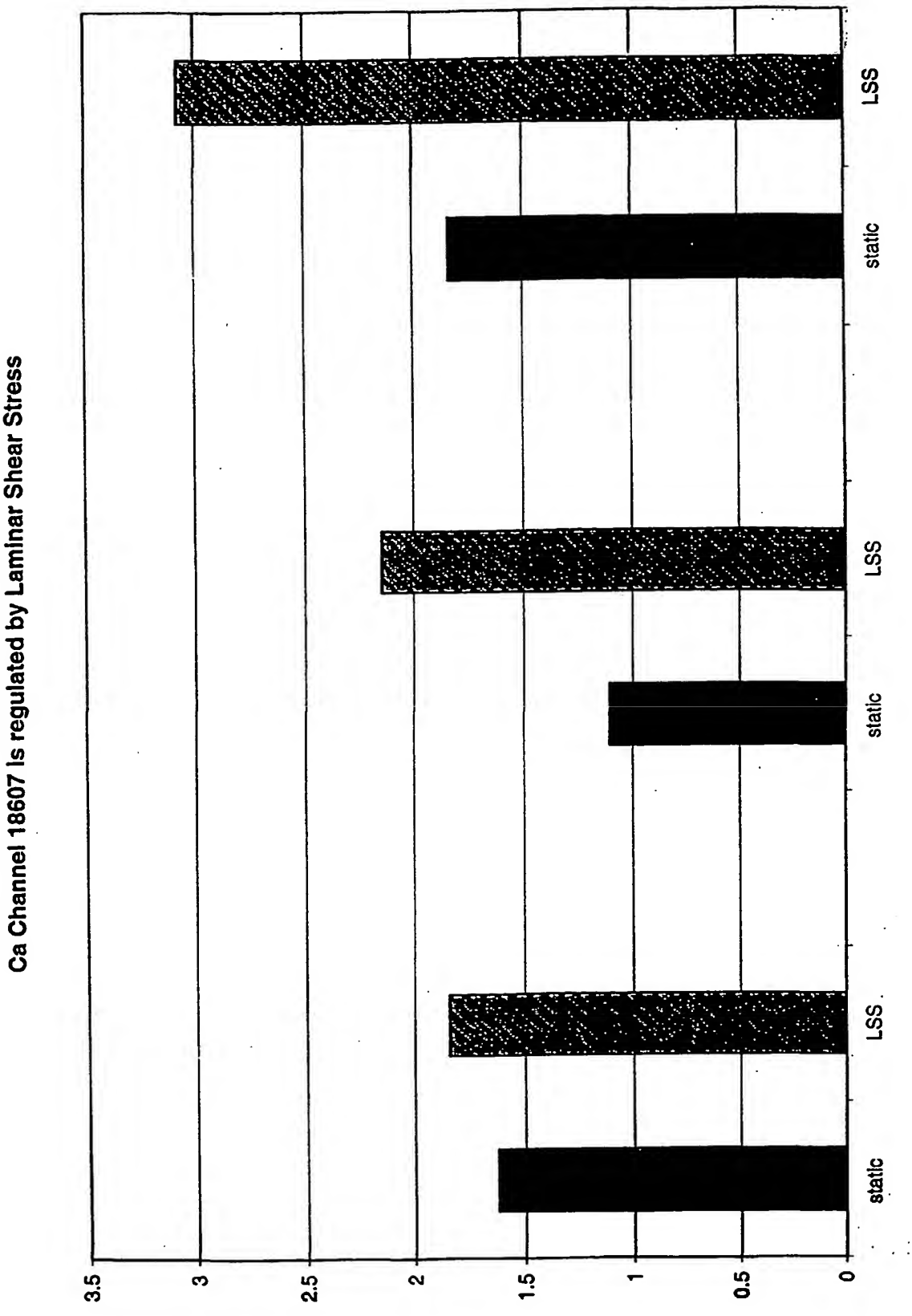

Ca Channel 18607 Is regulated by Laminar Shear Stress
3.5
3
2.5
2
1.5
1
0.5
0
LSS
static
LSS
static
LSS
static
LSS
static
static LSS
1.62183 1.840922
static LSS
1.106458 2.147504
static LSS
1.832452 3.076397

Figure 11

SEQUENCE LISTING

<110> MILLENNIUM PHARMACEUTICALS, INC.

<120> 18607, A Novel Human Calcium Channel

<130> MNI-097CP3PC

<140>
<141>

<150> 09/510,706
<151> 2000-02-22

<150> 09/634,669
<151> 2000-08-08

<150> 09/583,373
<151> 2000-05-21

<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 3900
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (138)..(3386)

<400> 1
cggcccatct ctctgggtct ctgtccctct ctctctgggt ctctgtcccc gtctctctgg 60

gtctcggtcc ccgtctctct gggtctctgt ccccgtctct ctgggtctct gtccccctcc 120

ctgtgtgccc cgctccc atg tgt cca cag ttc ctc cgg ctc tct gac cga 170
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg
1 5 10

acg gat cca gct gca gtt tat agt ctg gtc aca cgc aca tgg ggc ttc 218
Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe
15 20 25

cgt gcc ccg aac ctg gtg gtg tca gtg ctg ggg gga tcg ggg ggc ccc 266
Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro
30 35 40

gtc ctc cag acc tgg ctg cag gac ctg ctg cgt cgt ggg ctg gtg cgg 314
Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg
45 50 55

gct gcc cag agc aca gga gcc tgg att gtc act ggg ggt ctg cac acg 362
Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr
60 65 70 75

```
ggc atc ggc cgg cat gtt ggt gtg gct gta cgg gac cat cag atg gcc   410
Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala
                 80                  85                  90

agc act ggg ggc acc aag gtg gtg gcc atg ggt gtg gcc ccc tgg ggt   458
Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly
             95                 100                 105

gtg gtc cgg aat aga gac acc ctc atc aac ccc aag ggc tcg ttc cct   506
Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro
        110                 115                 120

gcg agg tac cgg tgg cgc ggt gac ccg gag gac ggg gtc cag ttt ccc   554
Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro
    125                 130                 135

ctg gac tac aac tac tcg gcc ttc ttc ctg gtg gac gac ggc aca cac   602
Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His
140                 145                 150                 155

ggc tgc ctg ggg ggc gag aac cgc ttc cgc ttg cgc ctg gag tcc tac   650
Gly Cys Leu Gly Gly Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr
                160                 165                 170

atc tca cag cag aag acg ggc gtg gga ggg act gga att gac atc cct   698
Ile Ser Gln Gln Lys Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro
            175                 180                 185

gtc ctg ctc ctc ctg att gat ggt gat gag aag atg ttg acg cga ata   746
Val Leu Leu Leu Leu Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile
        190                 195                 200

gag aac gcc acc cag gct cag ctc cca tgt ctc ctc gtg gct ggc tca   794
Glu Asn Ala Thr Gln Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser
    205                 210                 215

ggg gga gct gcg gac tgc ctg gcg gag acc ctg gaa gac act ctg gcc   842
Gly Gly Ala Ala Asp Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala
220                 225                 230                 235

cca ggg agt ggg gga gcc agg caa ggc gaa gcc cga gat cga atc agg   890
Pro Gly Ser Gly Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg
                240                 245                 250

cgt ttc ttt ccc aaa ggg gac ctt gag gtc ctg cag gcc cag gtg gag   938
Arg Phe Phe Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu
            255                 260                 265

agg att atg acc cgg aag gag ctc ctg aca gtc tat tct tct gag gat   986
Arg Ile Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp
        270                 275                 280

ggg tct gag gaa ttc gag acc ata gtt ttg aag gcc ctt gtg aag gcc   1034
Gly Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala
    285                 290                 295

tgt ggg agc tcg gag gcc tca gcc tac ctg gat gag ctg cgt ttg gct   1082
```

```
Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala
300                 305                 310                 315

gtg gct tgg aac cgc gtg gac att gca cag agt gaa ctc ttt cgg ggg   1130
Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly
                320                 325                 330

gac atc caa tgg cgg tcc ttc cat ctc gaa gct tcc ctc atg gac gcc   1178
Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala
            335                 340                 345

ctg ctg aat gac cgg cct gag ttc gtg cgc ttg ctc att tcc cac ggc   1226
Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly
        350                 355                 360

ctc agc ctg ggc cac ttc ctg acc ccg atg cgc ctg gcc caa ctc tac   1274
Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr
    365                 370                 375

agc gcg gcg ccc tcc aac tcg ctc atc cgc aac ctt ttg gac cag gcg   1322
Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala
380                 385                 390                 395

tcc cac agc gca ggc acc aaa gcc cca gcc cta aaa ggg gga gct gcg   1370
Ser His Ser Ala Gly Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala
                400                 405                 410

gag ctc cgg ccc cct gac gtg ggg cat gtg ctg agg atg ctg ctg ggg   1418
Glu Leu Arg Pro Pro Asp Val Gly His Val Leu Arg Met Leu Leu Gly
            415                 420                 425

aag atg tgc gcg ccg agg tac ccc tcc ggg ggc gcc tgg gac cct cac   1466
Lys Met Cys Ala Pro Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His
        430                 435                 440

cca ggc cag ggc ttc ggg gag agc atg tat ctg ctc tcg gac aag gcc   1514
Pro Gly Gln Gly Phe Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala
    445                 450                 455

acc tcg ccg ctc tcg ctg gat gct ggc ctc ggg cag gcc ccc tgg agc   1562
Thr Ser Pro Leu Ser Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser
460                 465                 470                 475

gac ctg ctt ctt tgg gca ctg ttg ctg aac agg gca cag atg gcc atg   1610
Asp Leu Leu Leu Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met
                480                 485                 490

tac ttc tgg gag atg ggt tcc aat gca gtt tcc tca gct ctt ggg gcc   1658
Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala
            495                 500                 505

tgt ttg ctg ctc cgg gtg atg gca cgc ctg gag cct gac gct gag gag   1706
Cys Leu Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu
        510                 515                 520

gca gca cgg agg aaa gac ctg gcg ttc aag ttt gag ggg atg ggc gtt   1754
Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
    525                 530                 535
```

```
gac ctc ttt ggc gag tgc tat cgc agc agt gag gtg agg gct gcc cgc   1802
Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg
540                 545                 550                 555

ctc ctc ctc cgt cgc tgc ccg ctc tgg ggg gat gcc act tgc ctc cag   1850
Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln
                560                 565                 570

ctg gcc atg caa gct gac gcc cgt gcc ttc ttt gcc cag gat ggg gta   1898
Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val
            575                 580                 585

cag tct ctg ctg aca cag aag tgg tgg gga gat atg gcc agc act aca   1946
Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr
        590                 595                 600

ccc atc tgg gcc ctg gtt ctc gcc ttc ttt tgc cct cca ctc atc tac   1994
Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr
    605                 610                 615

acc cgc ctc atc acc ttc agg aaa tca gaa gag gag ccc aca cgg gag   2042
Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu
620                 625                 630                 635

gag cta gag ttt gac atg gat agt gtc att aat ggg gaa ggg cct gtc   2090
Glu Leu Glu Phe Asp Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val
                640                 645                 650

ggg acg gcg gac cca gcc gag aag acg ccg ctg ggg gtc ccg cgc cag   2138
Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln
            655                 660                 665

tcg ggc cgt ccg ggt tgc tgc ggg ggc cgc tgc ggg ggg cgc cgg tgc   2186
Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys
        670                 675                 680

cta cgc cgc tgg ttc cac ttc tgg ggc gcg ccg gtg acc atc ttc atg   2234
Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met
    685                 690                 695

ggc aac gtg gtc agc tac ctg ctg ttc ctg ctg ctt ttc tcg cgg gtg   2282
Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val
700                 705                 710                 715

ctg ctc gtg gat ttc cag ccg gcg ccg ccc ggc tcc ctg gag ctg ctg   2330
Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu
                720                 725                 730

ctc tat ttc tgg gct ttc acg ctg ctg tgc gag gaa ctg cgc cag ggc   2378
Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly
            735                 740                 745

ctg agc gga ggc ggg ggc agc ctc gcc agc ggg ggc ccc ggg cct ggc   2426
Leu Ser Gly Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly
        750                 755                 760

cat gcc tca ctg agc cag cgc ctg cgc ctc tac ctc gcc gac agc tgg   2474
```

```
His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp
    765                 770                 775

aac cag tgc gac cta gtg gct ctc acc tgc ttc ctc ctg ggc gtg ggc  2522
Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly
780                 785                 790                 795

tgc cgg ctg acc ccg ggt ttg tac cac ctg ggc cgc act gtc ctc tgc  2570
Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys
                800                 805                 810

atc gac ttc atg gtt ttc acg gtg cgg ctg ctt cac atc ttc acg gtc  2618
Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val
            815                 820                 825

aac aaa cag ctg ggg ccc aag atc gtc atc gtg agc aag atg atg aag  2666
Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys
        830                 835                 840

gac gtg ttc ttc ttc ctc ttc ttc ctc ggc gtg tgg ctg gta gcc tat  2714
Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr
    845                 850                 855

ggc gtg gcc acg gag ggg ctc ctg agg cca cgg gac agt gac ttc cca  2762
Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro
860                 865                 870                 875

agt atc ctg cgc cgc gtc ttc tac cgt ccc tac ctg cag atc ttc ggg  2810
Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly
                880                 885                 890

cag att ccc cag gag gac atg gac gtg gcc ctc atg gag cac agc aac  2858
Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn
            895                 900                 905

tgc tcg tcg gag ccc ggc ttc tgg gca cac cct cct ggg gcc cag gcg  2906
Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala
        910                 915                 920

ggc acc tgc gtc tcc cag tat gcc aac tgg ctg gtg gtg ctg ctc ctc  2954
Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu
    925                 930                 935

gtc atc ttc ctg ctc gtg gcc aac atc ctg ctg gtc aac ttg ctc att  3002
Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile
940                 945                 950                 955

gcc atg ttc agt tac aca ttc ggc aaa gta cag ggc aac agc gat ctc  3050
Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu
                960                 965                 970

tac tgg aag gcg cag cgt tac cgc ctc atc cgg gaa ttc cac tct cgg  3098
Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg
            975                 980                 985

ccc gcg ctg gcc ccg ccc ttt atc gtc atc tcc cac ttg cgc ctc ctg  3146
Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu
        990                 995                1000
```

```
ctc agg caa ttg tgc agg cga ccc cgg agc ccc cag ccg tcc tcc ccg   3194
Leu Arg Gln Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro
   1005                1010                1015

gcc ctc gag cat ttc cgg gtt tac ctt tct aag gaa gcc gag cgg aag   3242
Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys
1020               1025                1030                1035

ctg cta acg tgg gaa tcg gtg cat aag gag aac ttt ctg ctg gca cgc   3290
Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg
               1040                1045                1050

gct agg gac aag cgg gag agc gac tcc gag cgt ctg aag cgc acg tcc   3338
Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser
           1055                1060                1065

cag aag gtg gac ttg gca ctg aaa cag ctg gga cac atc cgc gag tac   3386
Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr
       1070                1075                1080

gaacagcgcc tgaaagtgct ggagcgggag gtccagcagt gtagccgcgt cctggggtgg 3446

gtggccgagg ccctgagccg ctctgccttg ctgcccccag gtgggccgcc acccccctgac 3506

ctgcctgggt ccaaagactg agccctgctg gcggacttca aggagaagcc cccacagggg 3566

attttgctcc tagagtaagg ctcatctggg cctcggcccc cgcacctggt ggccttgtcc 3626

ttgaggtgag ccccatgtcc atctgggcca ctgtcaggac cacctttggg agtgtcatcc 3686

ttacaaacca cagcatgccc ggctcctccc agaaccagtc ccagcctggg aggatcaagg 3746

cctggatccc gggccgttat ccatctggag gctgcagggt ccttggggta acagggacca 3806

cagacccctc accactcaca gattcctcac actggggaaa taaagccatt tcagaggaaa 3866

aaaaaaaaaa aaaaarraaa aaaaaaaaag gcgg                              3900
```

<210> 2
<211> 1083
<212> PRT
<213> Homo sapiens

```
<400> 2
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg Thr Asp Pro Ala Ala
  1               5                  10                  15

Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe Arg Ala Pro Asn Leu
            20                  25                  30

Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro Val Leu Gln Thr Trp
        35                  40                  45

Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg Ala Ala Gln Ser Thr
    50                  55                  60
```

```
Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr Gly Ile Gly Arg His
 65                  70                  75                  80

Val Gly Val Ala Val Arg Asp His Gln Met Ala Ser Thr Gly Gly Thr
                 85                  90                  95

Lys Val Val Ala Met Gly Val Ala Pro Trp Gly Val Val Arg Asn Arg
            100                 105                 110

Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro Ala Arg Tyr Arg Trp
        115                 120                 125

Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr
    130                 135                 140

Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly
145                 150                 155                 160

Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys
                165                 170                 175

Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu
            180                 185                 190

Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln
        195                 200                 205

Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
    210                 215                 220

Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly
225                 230                 235                 240

Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys
                245                 250                 255

Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg
            260                 265                 270

Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe
        275                 280                 285

Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu
    290                 295                 300

Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg
305                 310                 315                 320

Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly Asp Ile Gln Trp Arg
                325                 330                 335

Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg
            340                 345                 350

Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His
        355                 360                 365

Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser
```

370 375 380

Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly
385 390 395 400

Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro
405 410 415

Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro
420 425 430

Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe
435 440 445

Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
450 455 460

Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu Trp
465 470 475 480

Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp Glu Met
485 490 495

Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Leu Arg
500 505 510

Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys
515 520 525

Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu
530 535 540

Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Leu Arg Arg
545 550 555 560

Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln Leu Ala Met Gln Ala
565 570 575

Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val Gln Ser Leu Leu Thr
580 585 590

Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu
595 600 605

Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr
610 615 620

Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp
625 630 635 640

Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro
645 650 655

Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly
660 665 670

Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe
675 680 685

```
His Phe Trp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
    690                 695                 700

Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp Phe
705                 710                 715                 720

Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe Trp Ala
                725                 730                 735

Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser Gly Gly Gly
            740                 745                 750

Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His Ala Ser Leu Ser
        755                 760                 765

Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp Asn Gln Cys Asp Leu
    770                 775                 780

Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys Arg Leu Thr Pro
785                 790                 795                 800

Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile Asp Phe Met Val
                805                 810                 815

Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln Leu Gly
            820                 825                 830

Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe Phe Phe
        835                 840                 845

Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala Thr Glu
    850                 855                 860

Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg
865                 870                 875                 880

Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu
                885                 890                 895

Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro
            900                 905                 910

Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser
        915                 920                 925

Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
    930                 935                 940

Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser Tyr
945                 950                 955                 960

Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys Ala Gln
                965                 970                 975

Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu Ala Pro
            980                 985                 990
```

```
Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln Leu Cys
        995                1000                1005

Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu His Phe
   1010                1015                1020

Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr Trp Glu
025                1030                1035                1040

Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp Lys Arg
               1045                1050                1055

Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val Asp Leu
           1060                1065                1070

Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr
       1075                1080


<210> 3
<211> 3387
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(3387)

<400> 3
atg tgt cca cag ttc ctc cgg ctc tct gac cga acg gat cca gct gca   48
Met Cys Pro Gln Phe Leu Arg Leu Ser Asp Arg Thr Asp Pro Ala Ala
  1               5                  10                  15

gtt tat agt ctg gtc aca cgc aca tgg ggc ttc cgt gcc ccg aac ctg   96
Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe Arg Ala Pro Asn Leu
             20                  25                  30

gtg gtg tca gtg ctg ggg gga tcg ggg ggc ccc gtc ctc cag acc tgg   144
Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro Val Leu Gln Thr Trp
         35                  40                  45

ctg cag gac ctg ctg cgt cgt ggg ctg gtg cgg gct gcc cag agc aca   192
Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg Ala Ala Gln Ser Thr
     50                  55                  60

gga gcc tgg att gtc act ggg ggt ctg cac acg ggc atc ggc cgg cat   240
Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr Gly Ile Gly Arg His
 65                  70                  75                  80

gtt ggt gtg gct gta cgg gac cat cag atg gcc agc act ggg ggc acc   288
Val Gly Val Ala Val Arg Asp His Gln Met Ala Ser Thr Gly Gly Thr
                 85                  90                  95

aag gtg gtg gcc atg ggt gtg gcc ccc tgg ggt gtg gtc cgg aat aga   336
Lys Val Val Ala Met Gly Val Ala Pro Trp Gly Val Val Arg Asn Arg
            100                 105                 110

gac acc ctc atc aac ccc aag ggc tcg ttc cct gcg agg tac cgg tgg   384
```

```
Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro Ala Arg Tyr Arg Trp
        115                 120                 125

cgc ggt gac ccg gag gac ggg gtc cag ttt ccc ctg gac tac aac tac   432
Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr
    130                 135                 140

tcg gcc ttc ttc ctg gtg gac gac ggc aca cac ggc tgc ctg ggg ggc   480
Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly
145                 150                 155                 160

gag aac cgc ttc cgc ttg cgc ctg gag tcc tac atc tca cag cag aag   528
Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys
                165                 170                 175

acg ggc gtg gga ggg act gga att gac atc cct gtc ctg ctc ctc ctg   576
Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu
            180                 185                 190

att gat ggt gat gag aag atg ttg acg cga ata gag aac gcc acc cag   624
Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln
        195                 200                 205

gct cag ctc cca tgt ctc ctc gtg gct ggc tca ggg gga gct gcg gac   672
Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
    210                 215                 220

tgc ctg gcg gag acc ctg gaa gac act ctg gcc cca ggg agt ggg gga   720
Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly
225                 230                 235                 240

gcc agg caa ggc gaa gcc cga gat cga atc agg cgt ttc ttt ccc aaa   768
Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys
                245                 250                 255

ggg gac ctt gag gtc ctg cag gcc cag gtg gag agg att atg acc cgg   816
Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg
            260                 265                 270

aag gag ctc ctg aca gtc tat tct tct gag gat ggg tct gag gaa ttc   864
Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe
        275                 280                 285

gag acc ata gtt ttg aag gcc ctt gtg aag gcc tgt ggg agc tcg gag   912
Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu
    290                 295                 300

gcc tca gcc tac ctg gat gag ctg cgt ttg gct gtg gct tgg aac cgc   960
Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg
305                 310                 315                 320

gtg gac att gca cag agt gaa ctc ttt cgg ggg gac atc caa tgg cgg  1008
Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly Asp Ile Gln Trp Arg
                325                 330                 335

tcc ttc cat ctc gaa gct tcc ctc atg gac gcc ctg ctg aat gac cgg  1056
Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg
            340                 345                 350
```

```
cct gag ttc gtg cgc ttg ctc att tcc cac ggc ctc agc ctg ggc cac   1104
Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His
        355                 360                 365

ttc ctg acc ccg atg cgc ctg gcc caa ctc tac agc gcg gcg ccc tcc   1152
Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser
    370                 375                 380

aac tcg ctc atc cgc aac ctt ttg gac cag gcg tcc cac agc gca ggc   1200
Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly
385                 390                 395                 400

acc aaa gcc cca gcc cta aaa ggg gga gct gcg gag ctc cgg ccc cct   1248
Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro
                405                 410                 415

gac gtg ggg cat gtg ctg agg atg ctg ctg ggg aag atg tgc gcg ccg   1296
Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro
            420                 425                 430

agg tac ccc tcc ggg ggc gcc tgg gac cct cac cca ggc cag ggc ttc   1344
Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe
        435                 440                 445

ggg gag agc atg tat ctg ctc tcg gac aag gcc acc tcg ccg ctc tcg   1392
Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
    450                 455                 460

ctg gat gct ggc ctc ggg cag gcc ccc tgg agc gac ctg ctt ctt tgg   1440
Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu Trp
465                 470                 475                 480

gca ctg ttg ctg aac agg gca cag atg gcc atg tac ttc tgg gag atg   1488
Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp Glu Met
                485                 490                 495

ggt tcc aat gca gtt tcc tca gct ctt ggg gcc tgt ttg ctg ctc cgg   1536
Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Leu Arg
            500                 505                 510

gtg atg gca cgc ctg gag cct gac gct gag gag gca gca cgg agg aaa   1584
Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys
        515                 520                 525

gac ctg gcg ttc aag ttt gag ggg atg ggc gtt gac ctc ttt ggc gag   1632
Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu
    530                 535                 540

tgc tat cgc agc agt gag gtg agg gct gcc cgc ctc ctc ctc cgt cgc   1680
Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Leu Arg Arg
545                 550                 555                 560

tgc ccg ctc tgg ggg gat gcc act tgc ctc cag ctg gcc atg caa gct   1728
Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln Leu Ala Met Gln Ala
                565                 570                 575

gac gcc cgt gcc ttc ttt gcc cag gat ggg gta cag tct ctg ctg aca   1776
```

```
Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val Gln Ser Leu Leu Thr
          580                 585                 590

cag aag tgg tgg gga gat atg gcc agc act aca ccc atc tgg gcc ctg   1824
Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu
        595                 600                 605

gtt ctc gcc ttc ttt tgc cct cca ctc atc tac acc cgc ctc atc acc   1872
Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr
    610                 615                 620

ttc agg aaa tca gaa gag gag ccc aca cgg gag gag cta gag ttt gac   1920
Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp
625                 630                 635                 640

atg gat agt gtc att aat ggg gaa ggg cct gtc ggg acg gcg gac cca   1968
Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro
                645                 650                 655

gcc gag aag acg ccg ctg ggg gtc ccg cgc cag tcg ggc cgt ccg ggt   2016
Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly
            660                 665                 670

tgc tgc ggg ggc cgc tgc ggg ggg cgc cgg tgc cta cgc cgc tgg ttc   2064
Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe
        675                 680                 685

cac ttc tgg ggc gcg ccg gtg acc atc ttc atg ggc aac gtg gtc agc   2112
His Phe Trp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
    690                 695                 700

tac ctg ctg ttc ctg ctg ctt ttc tcg cgg gtg ctg ctc gtg gat ttc   2160
Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp Phe
705                 710                 715                 720

cag ccg gcg ccg ccc ggc tcc ctg gag ctg ctg ctc tat ttc tgg gct   2208
Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe Trp Ala
                725                 730                 735

ttc acg ctg ctg tgc gag gaa ctg cgc cag ggc ctg agc gga ggc ggg   2256
Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser Gly Gly Gly
            740                 745                 750

ggc agc ctc gcc agc ggg ggc ccc ggg cct ggc cat gcc tca ctg agc   2304
Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His Ala Ser Leu Ser
        755                 760                 765

cag cgc ctg cgc ctc tac ctc gcc gac agc tgg aac cag tgc gac cta   2352
Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp Asn Gln Cys Asp Leu
    770                 775                 780

gtg gct ctc acc tgc ttc ctc ctg ggc gtg ggc tgc cgg ctg acc ccg   2400
Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys Arg Leu Thr Pro
785                 790                 795                 800

ggt ttg tac cac ctg ggc cgc act gtc ctc tgc atc gac ttc atg gtt   2448
Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile Asp Phe Met Val
                805                 810                 815
```

```
ttc acg gtg cgg ctg ctt cac atc ttc acg gtc aac aaa cag ctg ggg   2496
Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln Leu Gly
            820                 825                 830

ccc aag atc gtc atc gtg agc aag atg atg aag gac gtg ttc ttc ttc   2544
Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe Phe Phe
        835                 840                 845

ctc ttc ttc ctc ggc gtg tgg ctg gta gcc tat ggc gtg gcc acg gag   2592
Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala Thr Glu
    850                 855                 860

ggg ctc ctg agg cca cgg gac agt gac ttc cca agt atc ctg cgc cgc   2640
Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg
865                 870                 875                 880

gtc ttc tac cgt ccc tac ctg cag atc ttc ggg cag att ccc cag gag   2688
Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu
                885                 890                 895

gac atg gac gtg gcc ctc atg gag cac agc aac tgc tcg tcg gag ccc   2736
Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro
            900                 905                 910

ggc ttc tgg gca cac cct cct ggg gcc cag gcg ggc acc tgc gtc tcc   2784
Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser
        915                 920                 925

cag tat gcc aac tgg ctg gtg gtg ctg ctc ctc gtc atc ttc ctg ctc   2832
Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
    930                 935                 940

gtg gcc aac atc ctg ctg gtc aac ttg ctc att gcc atg ttc agt tac   2880
Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser Tyr
945                 950                 955                 960

aca ttc ggc aaa gta cag ggc aac agc gat ctc tac tgg aag gcg cag   2928
Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys Ala Gln
                965                 970                 975

cgt tac cgc ctc atc cgg gaa ttc cac tct cgg ccc gcg ctg gcc ccg   2976
Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu Ala Pro
            980                 985                 990

ccc ttt atc gtc atc tcc cac ttg cgc ctc ctg ctc agg caa ttg tgc   3024
Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln Leu Cys
        995                1000                1005

agg cga ccc cgg agc ccc cag ccg tcc tcc ccg gcc ctc gag cat ttc   3072
Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu His Phe
   1010                1015                1020

cgg gtt tac ctt tct aag gaa gcc gag cgg aag ctg cta acg tgg gaa   3120
Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr Trp Glu
1025                1030                1035                1040

tcg gtg cat aag gag aac ttt ctg ctg gca cgc gct agg gac aag cgg   3168
```

```
Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp Lys Arg
                1045                1050                1055

gag agc gac tcc gag cgt ctg aag cgc acg tcc cag aag gtg gac ttg   3216
Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val Asp Leu
            1060                1065                1070

gca ctg aaa cag ctg gga cac atc cgc gag tac gaa cag cgc ctg aaa   3264
Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr Glu Gln Arg Leu Lys
        1075                1080                1085

gtg ctg gag cgg gag gtc cag cag tgt agc cgc gtc ctg ggg tgg gtg   3312
Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg Val Leu Gly Trp Val
    1090                1095                1100

gcc gag gcc ctg agc cgc tct gcc ttg ctg ccc cca ggt ggg ccg cca   3360
Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro Pro Gly Gly Pro Pro
1105                1110                1115                1120

ccc cct gac ctg cct ggg tcc aaa gac                               3387
Pro Pro Asp Leu Pro Gly Ser Lys Asp
                1125
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 August 2001 (30.08.2001)

PCT

(10) International Publication Number
WO 01/62794 A3

(51) International Patent Classification[7]: C07K 14/705, C12N 15/11, 15/85, A61K 38/17, G01N 33/68, C12N 5/10, C07K 16/28

(21) International Application Number: PCT/US01/05529

(22) International Filing Date: 20 February 2001 (20.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/510,706	22 February 2000 (22.02.2000)	US
09/583,373	31 May 2000 (31.05.2000)	US
09/634,669	8 August 2000 (08.08.2000)	US

(71) Applicant *(for all designated States except US)*: **MILLENNIUM PHARMACEUTICALS, INC.** [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GLUCKSMANN, Maria, Alexandra** [AR/US]; 33 Summit Road, Lexington, MA 02173 (US). **CURTIS, Rory, A., J.** [US/US]; 31 Constitution Drive, Southborough, MA 01772 (US). **LORA, Jose, M.** [US/US]; 49 North Main Street, Natick, MA 01760 (US).

(74) Agents: **MANDRAGOURAS, Amy, E.**; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
2 May 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/62794 A3

(54) Title: 18607, A HUMAN CALCIUM CHANNEL

(57) Abstract: The invention provides isolated nucleic acids molecules, designated TLCC nucleic acid molecules, which encode novel TRP-like calcium channel molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TLCC nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TLCC gene has been introduced or disrupted. The invention still further provides isolated TLCC proteins, fusion proteins, antigenic peptides and anti-TLCC antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/05529

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/705 C12N15/11 C12N15/85 A61K38/17 G01N33/68
C12N5/10 C07K16/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, GENSEQ, WPI Data, PAJ, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16) claim 1B; figure 1	1-28
X	& DATABASE GENSEQ 'Online! AC No: AAV26656, 15 December 1998 (1998-12-15) COHEN M ET AL: "Human PS112 consensus DNA fragment from gene specific clones" Sequence with 98.9% identity with SEQ ID No:1 over 2395 nucleotides abstract	1-28

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 August 2001	24/08/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Nichogiannopoulou, A

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27) page 78 -page 79; claim 3	1-28
X	& DATABASE GENSEQ 'Online! AC No: AAV61200, 6 January 1999 (1999-01-06) DILLON DC ET AL: "Full length cDNA sequence of prostate tumour clone J1-17" Sequence with 99.8% identity with SEQ ID No:1 over 1522 nucleotides abstract	1-28
X	& DATABASE GENSEQ 'Online! AC No: AAW71868, 6 January 1999 (1999-01-06) DILLON DC ET AL: "Amino acid encoded by prostate tumour clone J1-17" Sequence with 100% identity with SEQ ID No:2 over 269 amino acids abstract	1-28

P,X	WO 00 40614 A (BETH ISRAEL HOSPITAL; SCHARENBERG ANDREW M (US)) 13 July 2000 (2000-07-13) the whole document	1-44
P,X	& DATABASE GENSEQ 'Online! AC No: AAY95436, 10 October 2000 (2000-10-10) SCHARENBERG AM : "Human calcium channel SOC-3/CRAC-2" Sequence with 100% identity with SEQ ID No:2 over 1079 amino acids abstract	1-44

A	ZHU X ET AL: "TRP, A NOVEL MAMMALIAN GENE FAMILY ESSENTIAL FOR AGONIST-ACTIVATED CAPACITATIVE CA2+ ENTRY" CELL, CELL PRESS, CAMBRIDGE, NA, US, vol. 85, no. 5, 1996, pages 661-671, XP000907242 ISSN: 0092-8674 the whole document	1-44

A	WES PAUL D ET AL: "TRPC1, a human homolog of a Drosophila store-operated channel" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 92, no. 21, 1995, pages 9652-9656, XP002138820 ISSN: 0027-8424 the whole document	1-44

	-/--	

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	SINKINS WILLIAM G ET AL: "Functional expression of TrpC1: A human homologue of the Drosophila Trp channel" BIOCHEMICAL JOURNAL, PORTLAND PRESS, LONDON, GB, vol. 331, no. 1, April 1998 (1998-04), pages 331-339, XP000864583 ISSN: 0264-6021 the whole document	1-44

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/05529

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9815657 A	16-04-1998	US 5919638 A	06-07-1999
		EP 0954599 A	10-11-1999
		US 6110675 A	29-08-2000
WO 9837093 A	27-08-1998	US 6261562 B	17-07-2001
		AU 731840 B	05-04-2001
		AU 6181898 A	09-09-1998
		CN 1252837 T	10-05-2000
		EP 1005546 A	07-06-2000
		HU 0002095 A	28-10-2000
		NO 994069 A	22-10-1999
		PL 335348 A	25-04-2000
		TR 9902053 T	21-04-2000
		US 6262245 B	17-07-2001
		ZA 9801585 A	04-09-1998
WO 0040614 A	13-07-2000	AU 2055600 A	24-07-2000